DOCUMENT RESUME

ED 137 333 TM 006 133

AUTHOR Milczarek, Gary J.
TITLE Instructional Development Formative Evaluation Design for EPEC.
INSTITUTION Ohio State Univ., Columbus. Evaluation Center.
PUB DATE Feb 73
NOTE 132p.; Not available in hard copy due to marginal legibility of original document

EDRS PRICE MF-$0.83 Plus Postage. HC Not Available from EDRS.
DESCRIPTORS Educational Change; *Educational Innovation; Educational Programs; Educational Researchers; *Evaluation; Evaluation Methods; *Formative Evaluation; *Instructional Systems; Professional Education; *Professional Personnel; *Skills
IDENTIFIERS *Evaluating the Process of Educational Change

ABSTRACT

In 1970 the Ohio State University Evaluation Center was funded by the Office of Education to design, operationalize, and implement a Model Training Project in Educational Evaluation. Of central importance was the generation of high quality, transportable instructional systems which were the responsibility of the Evaluation Center's Instructional Development (ID) Unit. This paper presents the ID Unit's design for the third developmental test of an instructional system called EPEC (Evaluating the Process of Educational Change). The EPEC instructional system provides participants opportunities to acquire knowledge, skills, and techniques for evaluating the installation of an educational innovation. The design calls for nine modules, approximately three hours each in length. Emphasis of the entire design is on participants practicing process evaluation skills. (RC)

INSTRUCTIONAL DEVELOPMENT

FORMATIVE EVALUATION DESIGN

for

EPEC

by

GARY J. MILCZAREK

THE OHIO STATE EVALUATION CENTER

Model Training Project for Educational Evaluation

The Ohio State University

College of Education

Columbus, Ohio

February, 1973

INSTRUCTIONAL DEVELOPMENT FORMATIVE EVALUATION DESIGN for EPEC

GARY J. MILCZAREK

THE OHIO STATE EVALUATION CENTER

ACKNOWLEDGEMENT

Two of my co-workers, Jack Sanders and Alan Nowakowski have contributed greatly to this evaluation design. Their valuable criticism has at least doubled its quality. I thank them both for many hours of helpful discussion and creative contribution.

Gary J. Milczarek

EPEC EVALUATION DESIGN

INTRODUCTION

In 1970 The Ohio State University Evaluation Center was funded by the Office of Education to design, operationalize and implement a Model Training Project (MTP) in Educational Evaluation. Of central importance to the project was the generation of high quality, transportable instructional systems. These instructional systems were the responsibility of the Evaluation Center's Instructional Development (ID) Unit. This paper presents the ID Unit's design for the third developmental test of an instructional system called EPEC, Evaluating the Process of Educational Change.

The Paper is organized according to the following topical outline:

1. Delineating Information Needs

- 1.1 THE EVALUATION SETTING
  - 1.1.1 Instructional Development Unit
  - 1.1.2 Description of EPEC
- 1.2 INFORMATION REQUIREMENTS
  - 1.2.1 Antecedents to the Evaluation
  - 1.2.2 The Decision Setting
  - 1.2.3 The Criterion Variables
  - 1.2.4 The Decision Rules
- 1.3 EVALUATION POLICIES
  - 1.3.1 Access to Data Sources
  - 1.3.2 Access to Evaluative Information
  - 1.3.3 Evaluation Responsibilities and Resources
  - 1.3.4 Reporting Audiences and Schedule
- 1.4 EVALUATION ASSUMPTIONS
  - 1.4.1 Sampling Assumptions
  - 1.4.2 Treatment Assumptions
  - 1.4.3 Measurement Assumptions
  - 1.4.4 Analysis Assumptions

2. Obtaining the Information

- 2.1 DATA COLLECTION
  - 2.1.1 Information Sources
  - 2.1.2 Instrumentation
  - 2.1.3 Data Collection Design

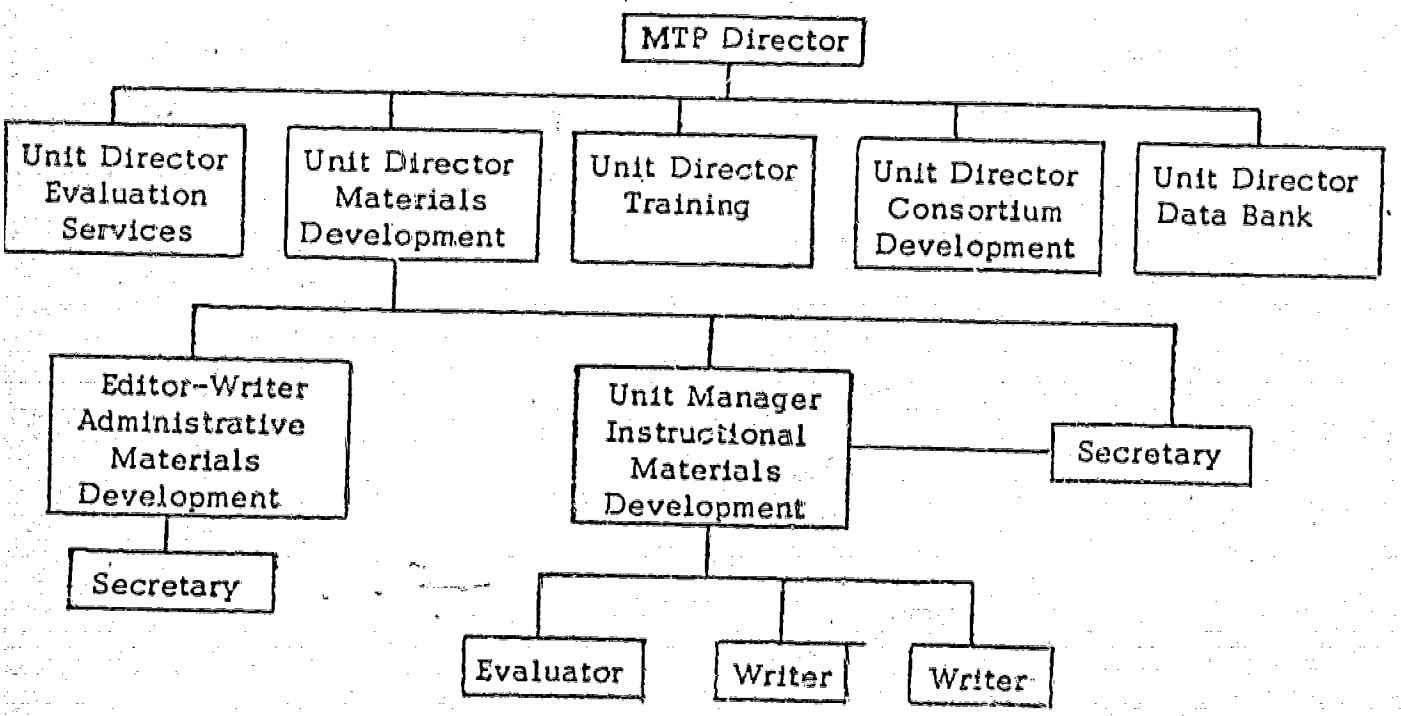
1. Delineation of Information Requirements

1.1 THE EVALUATION SETTING

1.1.1 Instructional Development Unit

The Instructional Development (ID) Unit is composed of the Unit Director and four Graduate Research Associates. One graduate student serves as manager of the team; another serves as evaluator. The Unit Director and remaining two students are writers for the team. Figure 1 presents an organizational chart depicting the ID Unit's relationship to the MTP.

Figure 1



Instructional Development is currently based on the Center Instructional Development (CID) Model presented in Appendix A. Since the development of EPEC was initiated before the development of the CID Model, some of the CID steps were not completed before the ID team began step 19.2, Developmental Test. In particular, the performance objectives were not written (step 7.2) and mastery items were not written (step 7.3) nor validated (step 7.6). EPEC has already undergone two developmental tests and the team is embarking on the third to begin November 1, 1972.

1.1.2 Description of EPEC

This package responds to the need identified by the OSU Model Training Project in Educational Evaluation for materials to train educational evaluators and evaluation-oriented educational leaders how to evaluate some aspects of the educational change process. The need was identified through a context evaluation performed by the adjunct professors, regular faculty and senior staff of the OSU Educational Development Faculty.

EPEC is a nine module instructional system consisting of 30 hours of instructional activities. The materials include an Instructor's Manual, Participant Manuals and one cassette tape.

A copy of the orientation to EPEC, explaining the purposes and describing the content, is included as Appendix B.

The system has three general purposes:
--to provide participants an experience in process evaluation
--to provide participants opportunities to try out some process evaluation skills
--to provide participants opportunities to identify the kinds of decisions serviced by a process evaluator

1.2 INFORMATION REQUIREMENTS

1.2.1 Antecedents to the Evaluation

The primary impetus for this study is that evaluation is called for in the CID Model (step 19.2.5) to provide the ID team with information to aid in revision of the instructional system. However, there is a secondary antecedent: since some of the CID Model development steps were not completed, and since considerable goal drift has been noted (see recorded documentation meeting of October 11, 1972) there is the possibility that the outcomes of EPEC in its current status may not be sufficiently isomorphic with the training goals of the MTP to warrant continued development. It is hoped that this evaluation will provide sufficient information about EPEC outcomes to determine whether continued revision is needed or warranted. (See minutes of the ID team meeting of October 25, 1972.)

1.2.2 The Decision Setting

The major decision question to be answered by this evaluation is: Do the EPEC learning activities need revision? Decisions regarding revision of EPEC are made by consensus among the ID team members. Evaluative information is thus to be directed to the team and preserved as a record for review by audiences specified under Evaluation Policies, below.

Evaluative information should be summarized for the ID team within two days of a module tryout. The summary must be appended with sufficient detail to permit in-depth analysis when desired by the team.

In order to determine the probable expected outcomes of EPEC a member of the ID team examined all the learning activities and listed the objectives that were apparently expected of each. Terminal objectives were separated from instrumental objectives and became the target of evaluation. Three kinds of information were then seen as needed:

1. the extent to which objectives are met in training,
2. the extent to which these objectives correspond to MTP training objectives and
3. the extent to which these objectives were valid and sufficient objectives of an instructional system on evaluating educational change.

The present developmental test is to focus on only the first two information needs with later evaluation studies focusing on the third.

To supplement information about objective attainment, an examination of unintended or un-explicated outcomes is also to be made.

1.2.3 The Criterion Variables

The basic decision alternatives include:

1. retain a particular activity
2. revise a particular activity
3. add a new activity

Criteria for retaining an activity are specified below. Failure to meet these criteria during the test indicates that the ID team should consider revising the activity or adding a new one.

For formative evaluation purposes, it is useful to examine a learning activity in terms of those characteristics or elements that are changeable.

The following characteristics were identified for the developmental test of EPEC.

1. Substantive Content: The ideas, skills, principles, concepts, etc. that are the instructional subject of the learning activity.
2. Purpose: A statement of intention concerning the substantive content of the learning activity.
3. Performanced-based Objectives: The operationalization of the learning activity's purpose--statements of what the participant will be able to do to demonstrate that the purpose of the learning activity is accomplished.

4. Media and Methods: The means by which the learning activity's substantive content is presented to the participants, e.g., a lecture, a role playing exercise, an observation schedule, a slide-tape show, or a theory paper.
5. Materials: The physical objects or apparatus used in the activity: the role play instructions, slides, tapes or printed theory papers.
6. Learning Environment Specifications: Any special instructions concerning the nature of the environment that must be established for the learning activity.
7. Entry Behaviors: (where applicable) General entry behaviors are specified for the package as a whole, however, some learning activities require special entry conditions to be met if the activity is to succeed.
8. Time Allotment: The amount of time allocated to an activity as specified in the activity's materials.
9. Feedback Procedures: The means of providing information concerning participant progress to the participant, the instructor, and the agency providing the training.
10. Placement: The chronological position of the activity within and across modules.

It is obvious that some criteria are more appropriate during the conceptualization and formulation stage of development than during the test stage. Therefore the criteria are divided into formulation and test criteria as follows.

Formulation Criteria: that set of criteria to be met while constructing the learning activities of the instructional system. Following the CID Model, these criteria are applied at several different points during the development process, e.g., Step 7: Define Performance Objectives or Step 18: Construct Prototype. A learning activity of the instructional system is not considered complete until the Formulation Criteria have been met for each of the learning activity's elements. Meeting these criteria results in a conceptually sound prototype needing only empirical validation. Formulation Criteria for each of A LEARNING ACTIVITY'S (LA) ELEMENTS INCLUDE:

1. Substantive Content of the LA:
   a. must be logically or theoretically relevant to the purpose of the module.
   b. must be validated by experts.

2. Purpose of the LA:
   a. must be consistant with the purpose of the module.
   b. must be accomplished as a necessary condition to meeting the purpose of the module. The conditional bases may be logical, empirical, or theoretical.

3. Performance-based Objectives of the LA:
   a. must state the learner.
   b. must state the behavior product.
   c. must state the cue or stimulus condition.
   d. must state the limits of acceptable response.
   e. must be logically related to the purpose of the LA so that if the objective is attained, it can be inferred that the purpose of the LA is partially accomplished.
   f. must as a set, sufficiently cover the substantive content of the activity.
   g. must take into account the limitations of reality.
   h. must be necessary to the LA purpose on logical or theoretical grounds.

4. Media and Methods of the LA:
   a. must be appropriate for
      1) the content,
      2) the target audience.
   b. must be consistant with
      1) learning theory,
      2) communication (a/v) theory.
   c. must be sufficient on logical or theoretical grounds to accomplish the LA purpose.
   d. must be sequenced in a consistant and logical manner.
   e. must be least costly and elaborate for the job.

5. Materials of the LA:
   a. must be editorially adequate.

6. Learning Environment Specifications of the LA:
   a. must be necessary and sufficient to accomplishing the activity's purpose on logical or theoretical grounds.

7. Entry Behaviors of the LA:
   a. must be necessary and sufficient to accomplishing the activity's purpose on logical or theoretical grounds.

8. Time Allotment of the LA:
   a. must be appropriate for the relative value of the activity.

9. Feedback Procedures of the LA:
   a. must be consistant with learning theory.
   b. must explicate a remediation procedure when applicable.

10. Placement of the LA:
    a. must be consistant with a model for the system.
    b. must be consistant with learning theory.

Test Criteria: that set of criteria which must be met during a developmental test in order to retain an LA. (An LA that does not meet these criteria must be considered for revision. If the LA is revised, the original formulation criteria must be re-applied.)

1. Substantive content of the LA:
    a. must not be provided by the instructional manager.
    b. should be perceived by participants as relevant.

2. Purpose of the LA: (none)

3. Performance-based objectives of the LA:
    a. must be attained.

4. Media and Methods of the LA:
    a. must be perceived as satisfactory by
        1) participants,
        2) the instructional manager.
    b. must be implemented as specified.

5. Materials of the LA:
    a. must be perceived by participants as
        1) ledgible
        2) convenient to use
        3) not overly expensive
        4) appealing, engaging
        5) understandable
    b. must be perceived by the instructional manager as appropriate
    c. must be durable enough to fulfill their function.

6. Learning Environment Specifications of the LA:
    a. must be perceived as implemented by the observer.
    b. must be perceived as necessary and sufficient by
        1) the observer,
        2) the instructor, and
        3) the participants.

7. Entry Behaviors of the LA:
    a. must be perceived as necessary, sufficient and met by:
        1) the observer, and
        2) the instructional manager.

b. may be objectively assessed when needed.

8. Time allotment of the LA:
   a. must correspond with the actual time consumed during the developmental test.
   b. must be perceived as appropriate by
      1) participants, and
      2) instructional manager

9. Feedback Procedures of the LA:
   a. must be perceived by participants as adequate and not disruptive.

10. Placement of the LA:
    a. must be perceived by participant as appropriate in terms of sequence and continuity.

In addition to information about specific learning activities it is also desirable to obtain some macro information about the system as a whole. Two additional test criteria are intended to provide this macro information:

11. The outcomes of EPEC must be isomorphic with the training goals of the MTP as perceived by
    a. participants
    b. EPEC developers

12. The participant perceived outcomes of EPEC must correspond to the outcomes intended by the development team.

Since all information needs can not be anticipated in this design, a summary evaluation activity to follow the last module is planned. The evaluator is to keep a log of important issues and questions that occur and remain unanswered during the test. Then the summary evaluation is to include a participant interview and other activities to provide the needed information.

Since this evaluation relates to a developmental test, ONLY THE TEST CRITERIA ARE TO PLAY A ROLE IN THE EVALUATION DESIGN. In Table 1 each of the Test Criteria is presented, followed by the criterion decision rule, source of the relevant data and name of the instrument designed to obtain the data.

1.2.4 The Decision Rules

If decision alternatives are to be weighed in light of specified criteria, the nature of the relationship between decision alternatives and criteria must be clear. This relationship may be made clear through the establishment of decision rules that describe what circumstances dictate the adoption of what decision alternatives. Decision rules are not difficult to determine when the decision involves choice among alternative strategies to a given

TABLE 1

Matrix of Information Requirements

Criteria	Decision Rule	Data Source	Instrument
1. Substantive content of the LA:			
a. must not be provided by the instructional manager	f provided content=0	Observer	Observation Sche
b. should be perceived by participants as relevant	positive response≥85%	Participant	PROBE
2. Purpose of the LA: (none)			
3. Performance-based Objectives of the LA:			
a. must be attained	Per Cent (x≥90%)≥90%	Participant	Performance Tests
4. Media and Methods of the LA:			
a. must be perceived as satisfactory by			
1. participants,	affirmative response≥85%	Participants	PROBE
2. the instructional manager,	affirmative response	Inst. Mgr.	Instructor PROBE
b. must be implemented as specified.	f deviations=0	Observer	Observation Sched
5. Materials of the LA:			
a. must be perceived by participants as acceptable*	affirmative response≥85%	Participant	PROBE
b. must be perceived by the instructional manager as appropriate.	affirmative response	Inst. Mgr.	Instructor PROBE
c. must be durable enough to fulfill their function	affirmative response≥85%	Participant	PROBE
6. Learning Environment Specifications of the LA:			
a. must be perceived as implemented by the observer	f departures=0	Observer	Observation Sched
b. must be perceived as necessary and sufficient by			
1. the observer,	f reported problems=0	Observer	Observation Sched
2. the instructor, and	affirmative response	Inst. Mgr.	Instructor PROBE
3. the participants.	affirmative response≥85%	Participant	PROBE
7. Entry Behaviors of the LA:			
a. must be perceived as necessary, sufficient and met			
1. the observer, and	f reported problem=0	Observer	Observation Sched
2. the instructional manager	affirmative response	Inst. Mgr.	Instructor PROBE

*See test criteria for a breakdown of this category.

TABLE 1 (cont.)

Criteria	Decision Rule	Data Source	Instrument
8. Time allotment of the LA:			
a. must correspond with the actual time consumed	f discrepancies=0	Observer	Observation Sched
b. must be perceived as appropriate by			
1. participants, and	affirmative response≥85%	Participant	PROBE
2. instructional manager	affirmative response	Inst. Mgr.	Instructor PROBE
9. Feedback Procedures of the LA:			
a. must be perceived by participants as adequate and not disruptive	affirmative response 85%	Participant	PROBE
10. Placement of the LA:			
a. must be perceived by participants as appropriate in terms of sequence and continuity	affirmative response 85%	Participant	Module Questionna
11. The outcomes of EPEC must be isomorphic with the training goals of the MTP as perceived by			
a. participants	undetermined	Participant	SAES
b. EPEC developers	undetermined	Developers	SAES
12. The participant perceived outcomes of EPEC must correspond to the outcomes intended by the development team	undetermined	Participants & Developers	Final Questionnaire

end. Once the relevant criteria have been identified, that strategy best meeting those criteria is the choice. However, in the instructional development process a different kind of decision setting often presents itself. The decision maker must choose between retaining or revising an LA based on such information as the LA's effectiveness in meeting objectives or its usefulness as perceived by participants. But, how well should objectives be met? How useful should training be perceived--by what proportion of participants? How does the decision maker avoid arbitrarily specifying the critical limits in a decision rule?

Various decision rules have been proposed for the instructional development decision to revise/retain an LA. An often used convention is the 90-90 decision rule to retain the unit if at least 90% of the participants meet at least 90% of the objectives. This decision rule is probably an operationalization of "most of the students should learn most of the material." A similar rule was proposed by FC Butler (1972) involving use of the standard deviation of the normal curve. After developing a sufficiently validated test 85% of a trained population should meet 100% of the objectives.

Decision rules such as these, stated strictly in terms of objective attainment, may be practical in relatively few situations. It is difficult to take such a decision rule seriously when repeated experience has shown the difficulty and cost in time and money to reach such a level of mastery. The decision maker must ask himself if he seriously intends to revise for mastery. If other variables enter into the decision, they should be included in the decision rule. The decision rule should be as functional as possible. But it is often difficult to explicate these variables.

These decision rules can also force the decision maker into a defensive posture, having to defend departures from the rule to those to whom he is accountable. The decision rule must be re-stated each time there is a departure and consequently the rule loses its value in providing for systematic consideration and selection of decision alternatives.

In order to avoid these problems and yet still avoid arbitration in stating critical limits it was decided that critical limits of the criterion variables should relate to the relative importance of the variable and secondly, that the decision rules should dictate not "revise" alternatives, but only "consider revision" alternatives. If evidence suggested a need for revision as dictated by the decision rule, then another set of practical criteria would enter in on the final decision. The effort to systematically determine and relate these practical criteria to the decision alternatives by means of decision rules was considered too great. These practical criteria may include:

1. Strength of the evidence suggesting revision.
2. Seriousness of the consequences of identified deficiencies.
3. Cost in time, effort and money to correct the system.
4. The ID team's confidence in the prognosis for the system's ailments.

During actual implementation of this phase of the evaluation design there is to be an examination of the kinds of variables that do enter into ID team discussions and the decisions made to determine the feasibility of systematically gathering data about these variables.

Decision Rules for affirming that criteria are met are listed in Table 1. Criteria that are met lead to the decision to retain a learning activity in its present form. Unmet criteria lead to the decision to consider revision.

In general, the critical limits for decision rules were set as follows:

For criteria that relate to participant attitudes toward training, the critical limits were set at 85% of the test population.

The 90-90 decision rule convention was adopted for attainment of objectives. Since the rule would only lead to a consideration of revision the developers could afford to make an error on the side of too strict a rule rather than pass over package weaknesses that might be fairly easy to remedy.

Criteria related to the correspondence between intended and actual procedures and time lines received zero frequency of deviations critical limits.

When considering participant feedback and recommendations it is realized that it is the quality of the remark, and not the number of people who confirm it only, that should be considered. Therefore it is necessary to prevent the decision rules from filtering out important information from the decision makers--as might happen if only those data not meeting "retain" criteria are displayed. To accomplish this, all open-ended remarks are summarized in the information displays for the decision makers.

1.3 EVALUATION POLICIES

1.3.1 Access to Data Sources

No important restrictions to data sources are foreseen. Generally, data sources include only the participants, instructors and observers of the developmental test.

1.3.2 Access to Evaluative Information

However, access to evaluative information will be restricted to the following personnel.

1. Assigned staff from the Evaluation Services Component
2. The MTP Director
3. USOE and NIE personnel assigned to monitor the MTP
4. The ID team

A final summary report of findings is to be made available to the public.

It is recognized that EPEC instructors and participants have a potential influence on revision decisions and might be considered an audience for evaluative information. However, since participants and instructors are the source of most of the data, they are not to be provided with evaluation results so as to avoid undue influence on their responses by knowledge of peer perceptions. It is expected that participants receive adequate feedback about their progress from the instructional materials.

1.3.3 Evaluation Responsibility and Resources

Evaluation responsibility is placed on the evaluator of the ID team. He may, however, delegate tasks and responsibilities to other ID team members as needed and appropriate. The resources allocated to this evaluation include one half time evaluator, a secretary and supplies as needed but are not expected to exceed 10% of the resources allocated to Instructional Development.

In order to avoid problems of co-optation and role conflict, it is understood that the evalutor's responsibility is limited to evaluation, that is, he does not engage in development of learning activities.

1.3.4 Reporting Audiences and Schedule

The ID team is provided with evaluative information within two days of each module's developmental test. A final report summarizing findings and each recommendation is presented within 30 days following the ninth module.

The evaluation design may be changed only by consensus of the ID team after examination of some of the evaluative data generated by the design.

1.4 EVALUATION ASSUMPTIONS

1.4.1 Sampling Assumptions

The target population for EPEC includes evaluation-oriented educators in pre-service or in-service training programs. However the evaluator does not have access to all possible members of the target population nor is it possible to set very adequate boundaries defining all possible members. There are no restrictions on the population except that college level learning skills are assumed. For this developmental test the sample will consist simply of the six graduate students who enrolled for EPEC Fall Quarter 1972. The intent is to sample their perceptions of EPEC and their performance to determine whether EPEC should be revised. Such a small sample represents a weakness in the design but yields more information than no test at all. When data are examined for their implications, the sample size should be kept in mind.

1.4.2 Treatment Assumptions

This design attempts to establish neither internal nor external validity of the treatment outcomes. The final report can not, therefore, attest with assurance that the outcomes are due to EPEC or are representative of outcomes that might be attained by other members of the population. This is a considerable weakness in the design and hopefully EPEC outcomes will be validated in further tests such as the Field Test. This developmental test is still expected to yield important information--whether or not the expected outcomes occur and whether or not actual outcomes are sufficiently isomorphic with the MTP's training goals.

1.4.3 Measurement Assumptions

There are three general classes of phenomena to be measured in this test. The first is participant and instructor affective reactions to specified variables of the learning activities. Essentially, it must be determined whether participants consider the activities to be satisfactory along the dimensions specified It is assumed that participants' responses to a scale of "satisfactoryness" will adequately reflect their affective reactions.

The second class of phenomena to be measured is the physical occurrence of specified events: trainer behaviors and procedural or temporal discrepancies. It is assumed that the frequency of these phenomena can best be measured by an observer using an observation schedule.

The third class of phenomena include participant performance in relation to EPEC's performance objectives. Perhaps the most questionable assumption is that these objectives are indeed a valid representation of the universe of process evaluation skills for change processes. Criterion referenced tests are used to measure objective attainment so the validity of the results hinge in part part on the validity of the inferences from correct responses to skill possession. Yet, do to a lack of time and resources, these inferences are not carefully examined and tested.

There are to be no formal tests of instrument reliability or objectivity except that performance tests will be scored independently by each of the ID team members and the results analyzed for agremment among scorers. The reporting parameters for this analysis will be a simple percent of scoring cases (across items and participants) in which there is perfect agreement among the five scores, and the percent of cases in which there is more than one dissenting score. A more sophisticated analysis is not warranted. The expected value for the percent of agreements is greater than 84%; for the percent of disagreements, less than 6%.

1.4.4 Analysis Assumptions

In general the information gathered is for comparison of results with a mastery or absolute standard. The standards are the criterion critical limits. The reporting parameters for these comparisons are frequencies and percentages. These analyses require no specific measurement assumptions. Comparison of outcome variables across activities and modules is in terms of the mean percent of respondents meeting criterion limits. No analyses requiring inferential statistics are expected.

2. Obtaining the Information

2.1 DATA COLLECTION

2.1.1 Information Sources

The sources of the information required for each of the evaluation criteria are presented in Table 1. The sample includes the six students who enrolled in EPEC for fall quarter plus two instructors, one for the first four modules and the other for the remaining five. Some information is to be provided by an observer. The sample used in this test is not necessarily representative of the target population of participants, instructors, or observers and no explicit inferences to these populations should be made. At the least, information about the sufficiency of EPEC is obtained for some members of the specified population.

2.1.2 Instrumentation

The instruments to be used in the design are displayed in Appendix D. Table 2 follows, with a list of these instruments.

TABLE 2

Instruments

Background Information Questionnaire
PROBE
Instructor PROBE
Module Questionnaire
Observation Schedule
Performance Tests
Final Questionnaire
Tape Recording of Sessions
Interview

2.1.3 Data Collection Design

Figure 2 portrays a general model of the evaluation design for a given module. All instruments are listed together with the individual who is to respond to the instrument and the administrator of the instrument. It is the observer's responsability to gather, organize and return the data gathered to the evalutor.

2.2 DATA ORGANIZATION

After a period of initial trial much of the data is to be stored and analyzed by computer. All data cards contain a common field including columns 1 through 10 storing the data coded in Table 3. This table also allocates card numbers to the various instruments. Instrument codes and column numbers are printed directly on the instruments.

2.2.1 The Unit of Organization

Subject by subject responses are retained as the unit of disaggregation.

2.2.2 Storage and Retrieval Requirements

The data is stored in raw form in a large loose-leaf binder under the care of the evaluator while analysis is underway. After the test, it is to be kept in a cabinet in the Resource Data Bank. Storage of data cards will be specified after the decision to begin use of computer data processing. The instruments and analysis forms are kept by the ID secretary.

2.3 DATA ANALYSIS

This section presents a strategy for obtaining needed information about EPEC revision from data collected in a module tryout.

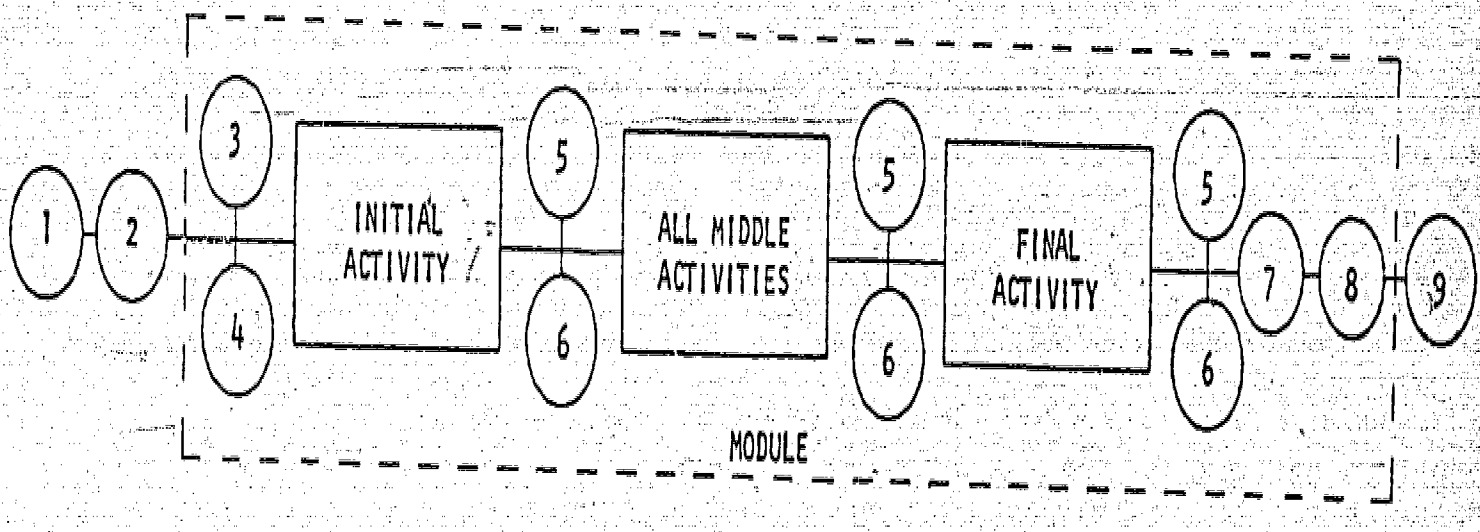
As indicated above, various criteria must be met by each LA if the LA is to stand unrevised. Specific decision rules indicate when the criteria are to be considered as satisfactorily met. After data are obtained the remaining task is to organize it in a way that shows whether these criteria have been met. To facilitate organization of the data a series of analysis forms are provided in Appendix E. Table 4 lists these analysis forms and shows their relation to the instruments and information sources presented above.

A brief overview of the analysis strategy is presented below. A more detailed presentation of the analysis instructions is provided in Appendix F.

The analysis procedure is basically a two stage process. The first stage involves completing the analysis forms which organize the data for the second stage. The second stage is an iterative process involving several steps. First, the

Figure 2

Flow-Chart of Evaluation Activities



	Instrument	Administrator	Recipiant
1 =	Background Information	Instructor	Participant & Instructor
2 =	PROBE General Instructions	Instructor	Participant
3 =	Observation Schedule	Observer	Module
4 =	Begin Tape Recording	Observer	Module
5 =	PROBE	Self	Participant
6 =	Instructor PROBE	Self	Instructor
7 =	Mastery Test	Instructor	Participant
8 =	Module Questionnaire	Instructor	Participant
9 =	Final Questionnaire	Instructor	Participant

TABLE 3

CODE SHEET

Common Field

Column	Variable and Code Range
01 - 02	Instructional System Number
03 - 04	Test Site Sequence Number
05	Instrument Number
	0 = Background Information Questionnaire
	1 = PROBE
	2 = Instructor PROBE
	3 = Performance Test
	4 = Module Questionnaire
	5 = Observational Schedule
	6 = Final Questionnaire
06	Module Number
07	Group Number
08 - 09	Participant ID Number
10	Open

TABLE 4

Information Provision
for
Module Report

Information Source	Instrument	Analysis Forms
Participant	PROBE	Table A-2 PROBE Results by Activity Table A-1 PROBE Summary Display A-1 PROBE Open Ended Remark
	Module Questionnaire	Table A-3 Questionnaire Results Display A-1 Questionnaire Open Ended Remarks
	Performance Test	Table A-4 Performance Test Results Table A-5 Performance Test Frequency Distribution
Instructor	Instructor Probe	Table B-1 Instructor Probe Ratings Display B-1 Instructor Comments
Observer	Observation Schedule	Table C-2 Agenda Item Data Table C-1 Activity Data Display C-1 Observation Comments
Evaluator		Table 1 Module Summary

criterion variable of interest to a particular element of an LA is identified. An example might be "performance-based objectives of the LA must be attained." Second, data relevant to this criterion variable are located on the analysis forms that were prepared in the first stage. Third, the results obtained for the criterion variable are noted in terms of the predetermined decision rule. If the criteria are met, the next criterion variable is considered. However, if criteria are not met, this result is noted in Table 1, the Module Summary. Entries in this table indicate that a particular element of the activity is not satisfactory and, through a coding system, refer the reader tosupporting information in the appendix of the report. Whenever an entry is made in Table 1, indicating a need for revision, a discussion of the results relating to the criterion variable under examination is developed in the body of the Module Report. This discussion considers all the information about the activity that is available, presents both the evidence suggesting revision and arguments for retaining the activity, offers explanations about the nature of the problem and recommends what action might be taken to improve the activity. The whole procedure is then repeated for the next criterion variable until all the elements of all the activities have been examined.

3. Providing the Information

3.1 DEFINITION OF THE REPORT AUDIENCE

The report audience for most of the evaluative information will include only the ID team. The Final Report is made available to all MTP component directors, NIE monitors, and other interested persons.

3.2 DESCRIPTION OF THE REPORTING MODE

The reports to the ID team consist of the completed analysis forms presented after each module of EPEC. Special staff meetings are established for the purpose of reviewing the information.

The Final Report is a printed paper condensing the major findings and conclusions of the evaluation study.

APPENDIX A

The Center Instructional Development Model

INSTRUCTIONAL DEVELOPMENT AT THE OHIO STATE UNIVERSITY EVALUATION CENTER

The Ohio State University Evaluation Center has evolved a systematic process for the development of educational products. This document describes this process, both to identify its concepts for Center personnel and to provide information for persons not affiliated with the Center.

August, 1972

Jack Sanders, Director
Instructional Materials Development

Paul Carlson, Research Associate

Jerry Adams, Research Associate

Kay Rofkahr, Media Specialist

Introduction

The Center Instructional Development (CID) Model presented here is not a new discovery. The CID Model has evolved from the many instructional development tasks performed by The Ohio State University Evaluation Center since its inception in 1965. Tasks, timelines, personnel, and products of the Center have varied over that seven years. But the dedication of the Center to the improvement of graduate education, generally, and the training of educational evaluators, specifically, has remained constant. The CID Model reflects the Center's dedication to instructional excellence as well as its past experiences in development. It builds on this dedication and experience, presenting a systematic process to guide the Center's future development efforts. Specifically, the CID Model is intended: a) to more efficiently fulfill the instructional development prescriptions of the Bargar "Model of Graduate Training" and the Bunda-Stufflebeam "Universe of Evaluation Competencies" substantive model; b) to delineate the processes followed by the Center in developing instructional products; c) to provide a system for internally and externally evaluating instructional development activities of the Center; d) to facilitate cooperative ventures with other developers and their agencies; e) to provide for students seeking evaluation/development assistantships an overview of the type of activities the Center's instructional development unit performs.

The CID Model is depicted in four phases: context phase, input phase, process phase, and product phase. Each of these phases are charted out in terms of task-flow and decision points. Next, task-flow is explicated along with appropriate criteria for decision points. Then, the entire array of CID Model processes is laid out for inspection. Finally, a design for validating the model is presented.

This model may or may not be linear. Under ideal conditions, where developers are not hampered by financial constraints, unrealistic timelines, personnel problems, and antiquated production facilities, the model may be consistantly followed in a linear fashion. But the Center has found that, even under "ideal conditions," strict linear use of the model is sometimes not feasible. For example, some Center work involves adapting or adopting products produced elsewhere. Naturally, the latter do not follow the process from beginning to end, but enter the process at an appropriate stage along the way. Additionally, some Center products may enter the model at more than one point simultaneously, e.g. a development team may be chosen before the product is through the context phase. Nevertheless, the model and processes presented here do reflect the standard course of development taken by most instructional products of the Center.

The CID Model is presented here as a prototype. It has been validated by external judges as conceptually adequate, but has not been through a complete pilot testing at the Center, nor has field testing established its generalizability. It is presented here as concrete evidence of the Center's

dedication to institutionalize a systematic approach to the development of instructional products.

Figure 1, following, charts the Context Phase of the CID Model.

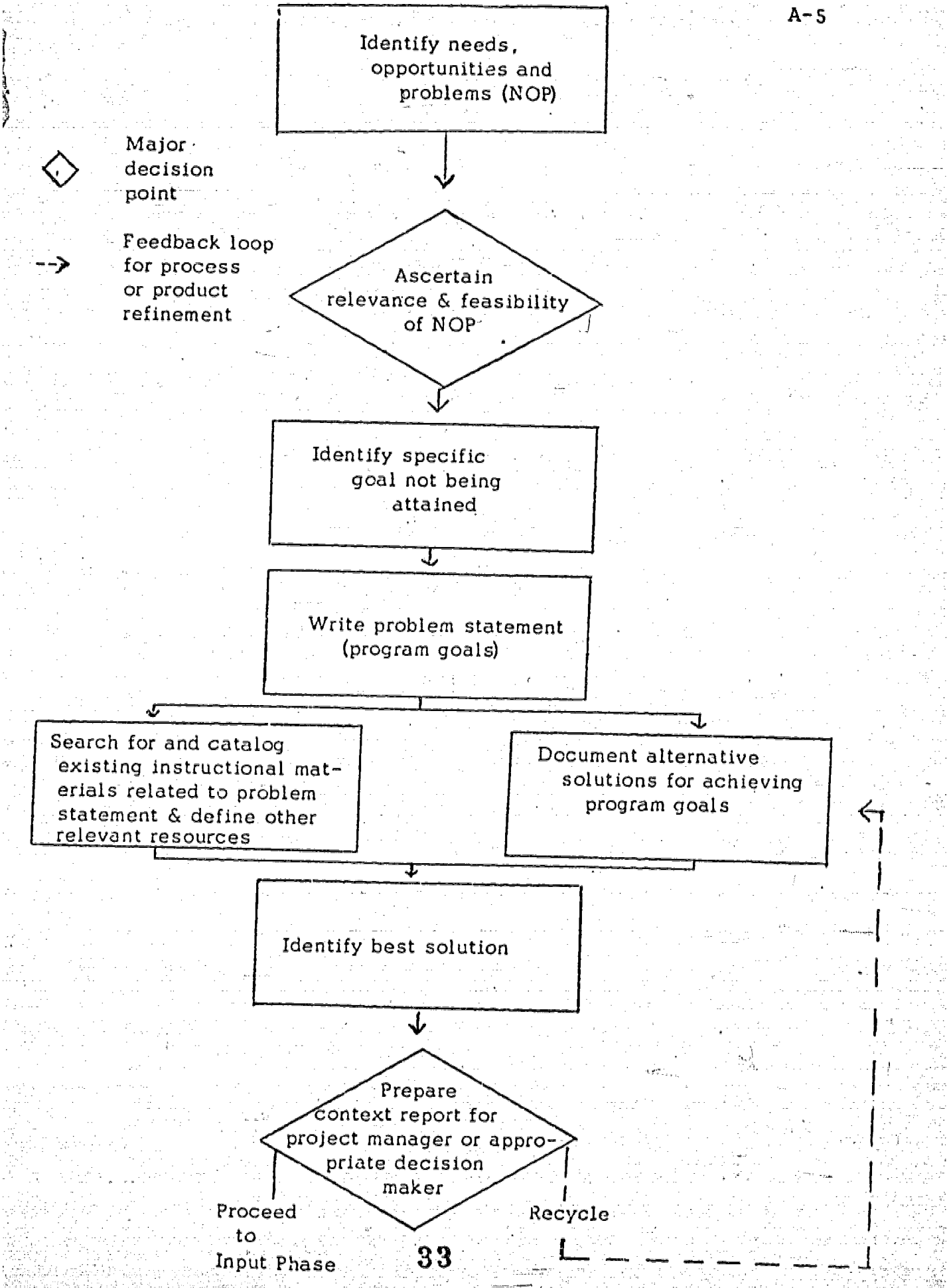


Figure 1: CONTEXT PHASE

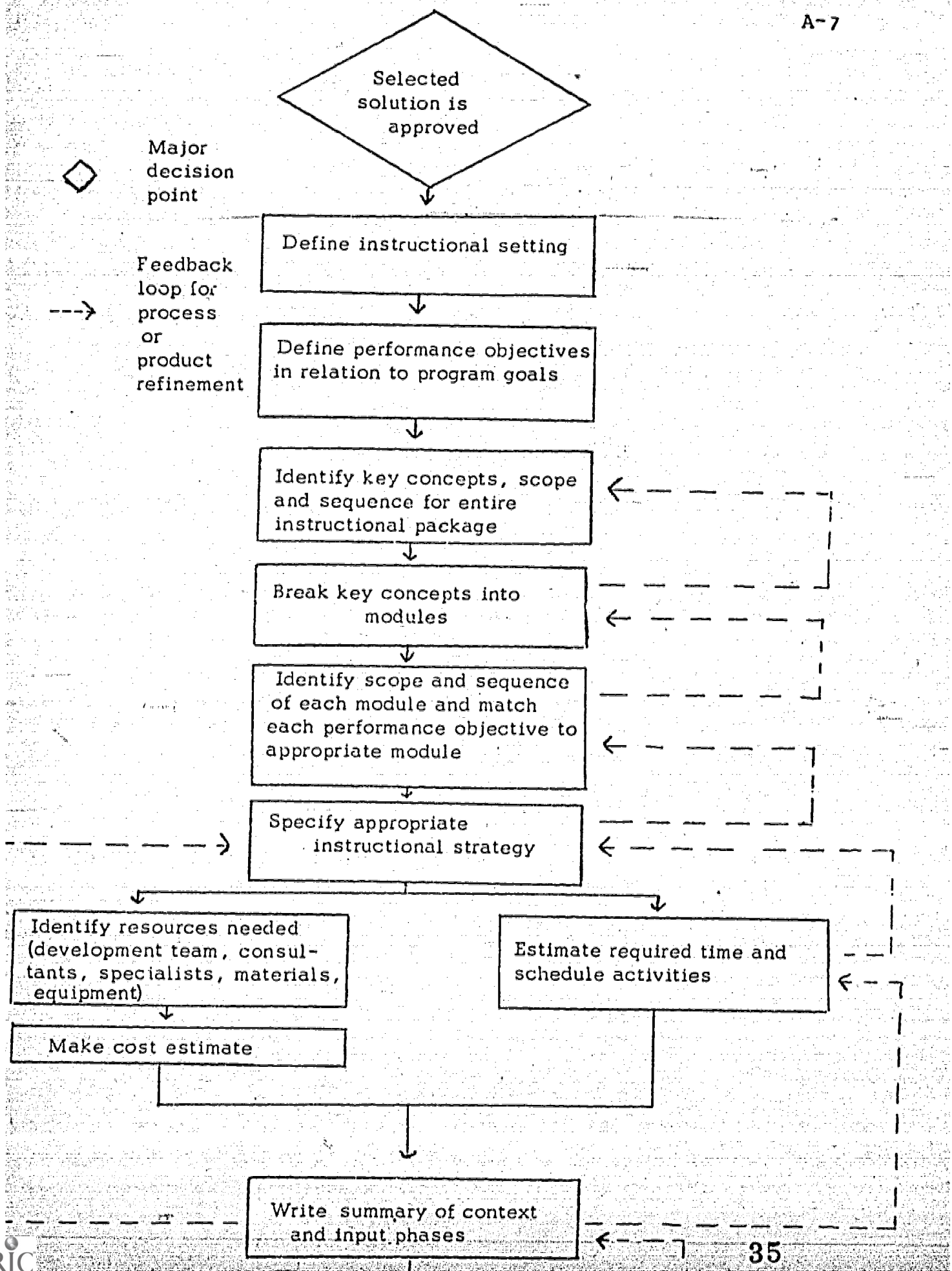
Context Phase

The Context Phase of the CID Model is concerned with delineating the evaluation training products that are needed, the unique opportunities which come to the Center for developing products, and the problems which prevent the Center from fulfilling identified needs or taking advantage of extant opportunities. The objective of the Context Phase is to prepare for the appropriate decision maker a report advocating start-up on development of the most needed instructional products.

Criteria for making decisions depicted in the Context Phase are:

1. That there is empirical support documenting the existence of specific needs.
2. That development opportunities available to the Center speak to documented needs.
3. That logical argument indicates there are no anticipated, insurmountable internal or external problems that would confront the developers in meeting the documented need.
4. That the proposed program's goals are not in conflict with the goals and subgoals of the Center.
5. That justification for the proposed "best solution" is empirically based.
6. That estimated cost-per-product is justified in terms of instructional need being served and number of potential users.

Figure 2, following, charts the Input Phase of the CID Model.


Selected solution is approved
Major decision point
Feedback loop for process or product refinement
Define instructional setting
Define performance objectives in relation to program goals
Identify key concepts, scope and sequence for entire instructional package
Break key concepts into modules
Identify scope and sequence of each module and match each performance objective to appropriate module
Specify appropriate instructional strategy
Identify resources needed (development team, consultants, specialists, materials, equipment)
Make cost estimate
Estimate required time and schedule activities
Write summary of context and input phases

Input Phase

The Input Phase of the CID Model is concerned with delineating the design of an instructional product which is to fulfill a high priority training need. The objective of the Input Phase is to prepare for the approval of the development team the design of a specific product and a detailed summary of resources necessary to implement the design.

Criteria for making decisions depicted in the Input Phase are:

1. That performance objectives and key concepts are as specific and clear as possible.
2. That key concepts are broken down into logically consistent modules.
3. That each module contains logically constant performance objectives
4. That the selected instructional strategy is consistent with the performance objectives.
5. That needed resources are available to development team.
6. That suggested time schedule can be adhered to by the development team.

Figure 3, following, charts the Process Phase of the CID Model.

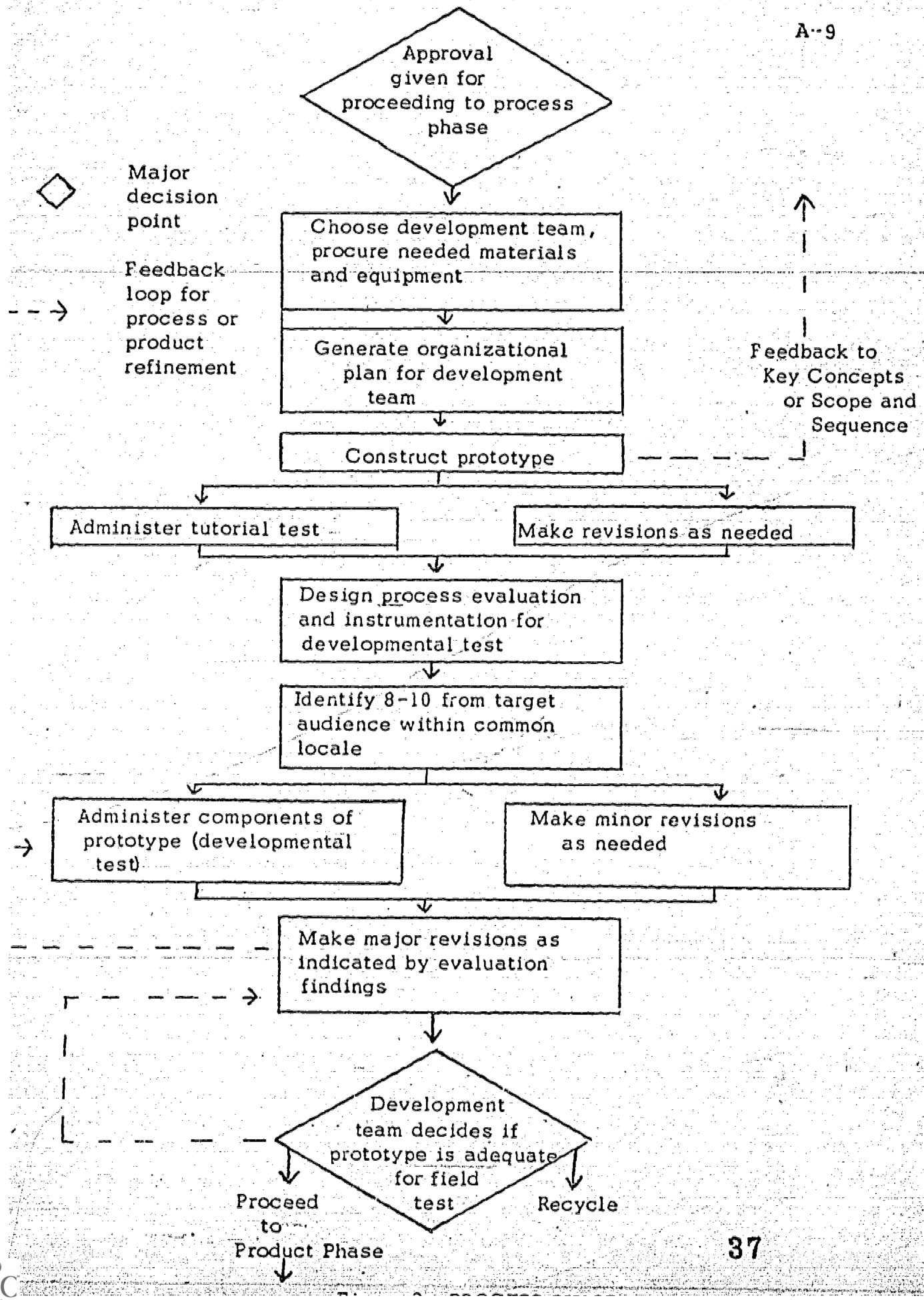

Approval given for proceeding to process phase
Major decision point
Feedback loop for process or product refinement
Choose development team, procure needed materials and equipment
Generate organizational plan for development team
Construct prototype
Feedback to Key Concepts or Scope and Sequence
Administer tutorial test
Make revisions as needed
Design process evaluation and instrumentation for developmental test
Identify 8-10 from target audience within common locale
Administer components of prototype (developmental test)
Make minor revisions as needed
Make major revisions as indicated by evaluation findings
Development team decides if prototype is adequate for field test
Proceed to Product Phase
Recycle

Figure 4: PRODUCT PHASE

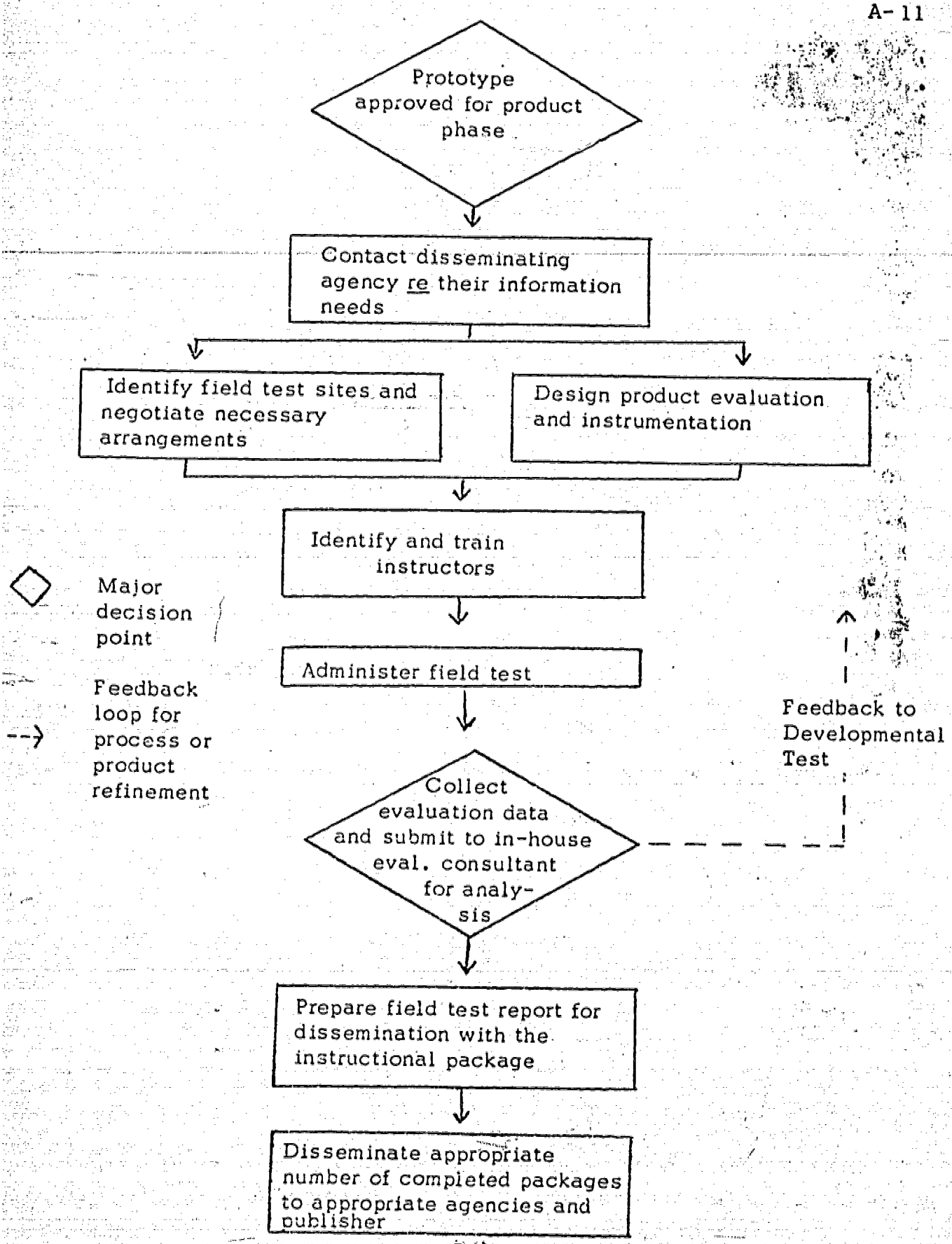
Process Phase

The Process Phase of the CID Model is concerned with building and testing a prototype instructional product. The objective of the Input Phase is to develop and refine the product so that it is satisfactory for field testing.

Criteria for making the decision that a product is ready for field test are:

1. Performance data indicates that product meets its performance objectives.
2. Adequacy of evaluation design, instrumentation, implementation, and report.
3. Adequacy of documentation of pilot test
4. Extent of necessary revisions after last developmental test.
5. Appropriate sample from target audience was used for field test.

Figure 4, following, charts the Product Phase of the CID Model.


Prototype approved for product phase
Contact disseminating agency re their information needs
Identify field test sites and negotiate necessary arrangements
Design product evaluation and instrumentation
Identify and train instructors
Administer field test
Collect evaluation data and submit to in-house eval. consultant for analy-sis
Prepare field test report for dissemination with the instructional package
Disseminate appropriate number of completed packages to appropriate agencies and publisher
Major decision point
Feedback loop for process or product refinement
Feedback to Developmental Test

Product Phase

The Product Phase of the CID Model is concerned with the objective field testing (summative evaluation) of an instructional product. The objective of the Product Phase is to empirically establish that the instructional product does meet the need it was intended to meet and hence is ready for dissemination.

Criteria for making the decision that development of a product has been successfully completed are:

1. The benefits of applying the product outweigh the costs of adopting it.
2. External evaluator's findings certify objectivity of field test.
3. Field test data indicates that performance objectives were attained.
4. Field test data indicates that instructor training was adequate.
5. Field test data indicates positive affective response of instructors and participants to the instructional product.
6. Field test data indicates the product is politically and morally viable for use of target audience.
7. Field test findings indicate the product (with minor revisions) is ready for dissemination.

Thus far, this document has considered the major activities, decision points, and criteria of the CID Model. Subsequently, a complete delineation of major and minor tasks associated with the Center's instructional development is presented.

1.0 Define Need/Opportunities/Problems (NOP)
- 1.1 Define source of NOP
  - 1.1.1 Students in evaluation training
  - 1.1.2 Curriculum Task Force
  - 1.1.3 Universe Data
  - 1.1.4 Training needs-analysis of Consortium
  - 1.1.5 Associated OSU Faculty, Dept., or Center
  - 1.1.6 RFP
  - 1.1.7 Oregon Data
  - 1.1.8 Other
- 1.2 Write NOP as problem statement (program goals)
  - 1.2.1 Identify specific goal not being attained
- 1.3 Validate problem statement with other sources
  - 1.3.1 Students in Evaluation training (principal source for validation)
  - 1.3.2 affiliated evaluation agencies
  - 1.3.3 Universe Data
  - 1.3.4 Other labs and centers
- 1.4 Begin log of development effort

2.0 Define Available Resources
- 2.1 Search for and catalog existing instructional materials related to problem statement
  - 2.1.1 PRDB
  - 2.1.2 NCERD "Hot Line"
  - 2.1.3 CEDaR
  - 2.1.4 TAP
  - 2.1.5 DID-AE(contacts
  - 2.1.6 Other
- 2.2 Content experts
  - 2.2.1 in-house
  - 2.2.2 other
- 2.3 Instructional Developers
  - 2.3.1 in-house
  - 2.3.2 other
- 2.4 Evaluators
  - 2.4.1 in-house
  - 2.4.2 other
- 2.5 Consultants, i.e., media specialists
  - 2.5.1 in-house
  - 2.5.2 other
- 2.6 Finances
  - 2.6.1 from existing development funds
  - 2.6.2 from new funds
  - 2.6.3 from other sources

2.7 Production facilities
2.8 Probable sites for tutorial, developmental and field tests

3.0 Document Alternative Solutions for Achieving Program Goals

4.0 Identify Best Solution

5.0 Prepare Context Report for Decision Maker

6.0 Define Instructional Setting
6.1 Target Audience
6.1.1 Graduate Students
6.1.2 Working evaluators
6.1.3 Beginning paraprofessionals
6.1.4 Others
6.2 Physical facilities for target audience
6.2.1 Space
6.2.2 Hardware
6.2.3 Software
6.2.4 Library
6.2.5 Other
6.3 Instructors, if appropriate
6.4 Community resources

7.0 Define Performance Objectives in Relation to Program Goals
7.1 Research existing banks of objectives
7.2 Write finalized performance objectives
7.3 Write mastery items
7.4 Write needed entry behaviors
7.5 Select sample from target audience
7.6 Validate mastery items
7.7 Select principles from learning theory and the psychology of learning appropriate to materials being developed
7.8 Inform associated agencies of specific development project via the Center Newsletter

8.0 Identify Key Concepts, Scope and Sequence for Entire Instructional Package

9.0 Break Key Concepts Into Modules

10.0 Identify Scope and Sequence of Each Module and Match Each Performance Objective to Appropriate Module

11.0 Specify Appropriate Instructional Strategy
- 11.1 Determine appropriate class size
  - 11.1.1 Large group (10+)
  - 11.1.2 Small group (2-9)
  - 11.1.3 Individuals
- 11.2 Determine instructional format
  - 11.2.1 Instructor required
    - 11.2.1.1 lecture
    - 11.2.1.2 team-taught
    - 11.2.1.3 symposium
    - 11.2.1.4 audio-tutorial
    - 11.2.1.5 other
  - 11.2.2 Instructor not required
    - 11.2.2.1 IPI
    - 11.2.2.2 CAI (or variants thereof)
    - 11.2.2.3 other
- 11.3 Determine instructional media
  - 11.3.1 print
    - 11.3.1.1 simulation
    - 11.3.1.2 games
    - 11.3.1.3 case studies
    - 11.3.1.4 other
  - 11.3.2 non-print
    - 11.3.2.1 CAI (or variants thereof)
    - 11.3.2.2 audio-tape
    - 11.3.2.3 video-tape
    - 11.3.2.4 slide-tape
    - 11.3.2.5 transparencies
    - 11.3.2.6 video-cassette
    - 11.3.2.7 other
- 11.4 Bibliography

12.0 Identify Resources Needed (development team, consultants, specialists, materials, equipment)
- 12.1 Make cost estimate

13.0 Estimate Time Required and Schedule Activities

14.0 Write Summary of Context and Input Phases

15.0 Submit Summary to Development Unit for Approval

16.0 Choose Development Team(s); Procure Needed Materials and Equipment
- 16.1 Division of labor
  - 16.1.1 content
  - 16.1.2 evaluation
  - 16.1.3 media
  - 16.1.4 instructional design
  - 16.1.5 editor
  - 16.1.6 programmed instruction
  - 16.1.7 CAI man
  - 16.1.8 other

17.0 Generate Organizational Plan for Development Team(s) (develop PERT chart with time-lines)
- 17.1 Scheduling of development team members
- 17.2 Develop communication network between and within teams; between/among administrators, developers, evaluators
- 17.3 Develop mechanism for resolving team conflict/deficiencies
- 17.4 Inform associated agencies of development timelines in the Center Newsletter

18.0 Construct Prototype
- 18.1 Review all previous decisions made
- 18.2 Review problem statement and objectives
- 18.3 Secure developmental copyright from O.E.

19.0 Test and analyze prototype
- 19.1 Tutorial test
  - 19.1.1 Identify student from target population
  - 19.1.2 Administer components of the prototype
  - 19.1.3 Revise
- 19.2 Developmental Test
  - 19.2.1 Develop evaluation criteria for test
  - 19.2.2 Identify group of 8-10 in common locale from target audience
  - 19.2.3 Develop formative evaluation methodology (and instruments, if appropriate)
  - 19.2.4 Administer components of the prototype (developmental test)
  - 19.2.5 Revise package (someone other than the author)
  - 19.2.6 Make adequate number of copies of the package
- 19.3 Field Test
  - 19.3.1 Contact disseminating agency re their information needs

19.3.2 Develop evaluation criteria
19.3.3 Identify five field test sites each with 8-10 students and negotiate necessary arrangements
19.3.4 Develop summative evaluation methodology and instruments
19.3.5 Identify and train instructors, if appropriate
19.3.5.1 on-site
19.3.5.2 at Center
19.3.6 Administer package to students
19.3.7 Collect summative data
19.3.8 Analyze summative data
19.3.9 Prepare field test report for dissemination with the package

20.0 Implement and Diffuse
20.1 Send five completed packages to OE
20.2 Send five completed packages to PRDB
20.3 Send description of package to associated agencies, TAP, and others
20.4 Select diffusion strategy
20.4.1 Press coverage
20.4.2 Informational letters to CEDaR, TAP, and R&D centers
20.4.3 Samples to leaders in the field
20.4.4 Other
20.5 Assist PRDB with diffusion until termination of project

Part Three: Testing the CID Model

Testing the usefulness of the model will involve a) securing professional judgments as to its internal consistency and face validity; b) securing the approval of the Center staff to implement the model; c) systematically documenting the appropriateness and usefulness of the model after installation.

Experts associated with the Center will be requested to evaluate the CID model in terms of its logic, completeness and communicability. This was done during the July, 1972 meeting of adjunct professors from associated evaluation agencies.

Following revision of the model, based on the comments of the experts, the Center staff agreed to an experimental implementation of the model with review of its effects after one year. Implementing this model in any agency requires a) agency consensus on the appropriateness of the model as a guide for agency instructional development activity; b) commitment on the part of the agency developers to the operationalization or adoption or adaptation and utilization of a design for testing the model; c) comparison of the quality of products produced vis-a-vis the model with that of previously produced products; d) commitment of sufficient resources to instructional development.

After installation of the CID model, the documentation of its usefulness will begin using three instruments: a check list, a log, and a chart. An explanation of each instrument follows.

The check list will contain items from the model that are checked off (and dated) as they are used or pursued. Figure 5 depicts the format of the CID check list.

DATE	CHECK AS COMPLETED	
_____	_____	1.0
_____	_____	1.1
_____	_____	1.1.1
_____	_____	1.1.2

Figure 5: CID CHECK LIST

The instructional development unit will use the Time and Effort Log as shown in Figure 6. However, in order to indicate more specifically what activities are pursued, the following code will be used:

A - "Evaluating the Educational Change Process" package
B - "Survey of Evaluation" package
C - Searching for other packages
D - Project coordination (e.g. staff meetings)
E - Center coordination (e.g. Center administration meetings)
F - Faculty/College coordination (e.g. committee meetings)

Return completed log to Ken Eye, 215 Oxley, at the Wednesday MTP Meeting the week following the period reported.

TIME & EFFORT LOG

Component Director [ ]

Name: _____ Position: Research Associate [ ]

Other _____

Date: The week beginning Monday, _____, thru Sunday, _____

	Writing Reports	Reading Documents	Person to Person Consultation	External Consultation	Meetings	Advisement	Editing, Proofing	Developing, eg Conceptualizing, Field Testing, etc.	General Clerical eg, Collating	Correspondence	Phone	MTP Teaching	Administrative Activities		
MON															
TUE															
WED															
THU															
FRI															
SAT															
SUN															

1. Use hours or part thereof of time allocated in appropriate cells.
2. Specify other activities in "blank" cells.

Products or Processes (Tasks) Completed:

Wrap Up:

Comments:

The appropriate letter will be written in the cell along with the amount of time spent. Thus, if the developer has spent two and one-half hours "searching for other packages," the code will be "2.5C". The search was probably conducted by "reading documents," "person to person consultation," "correspondence," or "phone"--all of which are headings above a column of cells. The appropriate cell is chosen and "2.5C" is written in.

If, however, an instructional package is the subject of the activity, then the identification numbers from the CID model can be used in combination with the above instructions. For example, if the developer spends four hours in preparing the field test report for a certain instructional system the code written in the log is: "4A 10.1.2". The appropriate cell, in this case, would be in the column under "Writing Reports."

A wall chart listing products (and dates) will be maintained for the entire project. The products could be added to the list daily and include "mini-products," such as memos, individual transparencies, and mockups. A format for the chart is depicted in Figure 7.

Product	Date	Product	Date	Product	Date
1		5		9	
2		6		10	
3		7		11	
4		8		12	

Figure 7: CID PRODUCTS CHART

Finally, during the instructional development of each product, additional criteria will be generated to service the decisions identified in the four phase-charts. The criteria will be subjected to comparative analysis in order to discover common and variant criteria which can be applied appropriately to different types of instructions products. For example, it may be possible to identify sets of criteria that will be appropriate to the development of two-day workshops as opposed to 30-hour instructional system packages.

year after it is installed, the CID model's usefulness should be a matter of record. If its positive utility is empirically established, the documentation effort would be refocused to address questions such as:

1) Would other representations of the general model in the form of a PERT chart, a critical path network, a time pie, or a moving wheel render it more communicable?

2) Can the components of the model be displayed or coded in more appealing ways, such as color coding, problem cards, or gameboards?

APPENDIX B: ORIENTATION TO EPEC

ORIENTATION TO E

Purposes and Objectives of EPEC

The initials "EPEC" stand for Evaluating the Process of Educational Change. There are three purposes to the EPEC system. One is to help you identify the kinds of decisions an evaluator services. The second is to provide you with an experience in evaluating the installation of a particular educational change or innovation. The third is to try out some specific process evaluation skills.

The evaluation profession is one of the newest specialties in the field of education. Its growth is attributable to a variety of circumstances, not the least of which are the needs of the educational consumer. Taxpayers, parents and students have, in recent years, become more insistent that educators be accountable for their decisions. These consumers want to know how curriculum, transportation, supervision and other innovations are going to benefit them. Thus, new university- and field-based training programs have been created to prepare skilled professionals who can provide educational consumers with the information they require.

Professional evaluators must be able to provide many different kinds of information to their clients. EPEC is concerned with some of the skills and knowledges evaluators need in order to provide a particular kind of information -- information concerning the degree to which some innovation is being successfully installed. These skills and knowledges are communicated by involving you first in a simulated evaluation task. You will experience that task first hand in a controlled setting. During the task you will play out the roles of

both the advocates and consumers of a particular educational change attempt. By playing these roles in a controlled setting, you will become sensitive to the significant influence advocates and consumers have on educational innovation. As a consequence of this sensitivity, you should be better able to evaluate the process of installing educational innovations. Then, by negotiated contract with your instructor, you will apply your learnings from the simulated task to evaluating the installation of an innovation in a real educational setting. But now we are getting ahead of ourselves.

The EPEC instructional system provides opportunities for participants to develop the following competencies:

—describing variables associated with the advocate of an educational innovation.

—describing variables associated with the consumer of an educational innovation.

—describing variables associated with the organization installing an educational innovation.

—describing variables associated with the educational innovation selected for installation.

—describing the kinds of transactions which occur between advocates and consumers during the installation of an educational innovation in an organization.

—applying the force field diagnostic technique.

—identifying the kinds of process decisions an evaluator services.

—developing/selecting and using process evaluation instruments.

—analyzing and reporting major results in data collected.

—utilizing concepts and skills of giving and receiving feedback.

—identifying and evaluating some group process skills as they relate to installing an innovation.

—writing a contingency contract to evaluate the real installation of an educational innovation.

Nature of EPEC

The EPEC instructional system provides participants opportunities to acquire knowledge, skills and techniques for evaluating the installation of an educational innovation. It looks at only one ridge of that living mountain called "educational evaluation." The design calls for nine modules, approximately three hours each in length. Emphasis of the entire design is on participants practicing process evaluation skills. Continuous active participation is demanded by using a simulated situation in which the participant assists a fictitious project director to install Computer Assisted Instruction (CAI) in a fictitious school using the skills in EPEC. In groups of three and six, participants use group process techniques such as role playing and brainstorming to apply and reinforce their learning. Prepared readings assigned at the end of several modules provide subject matter for subsequent modules. The carrying out of the individual contingency contracts is the only other out-of-class work required.

EPEC is concerned with part of the problem solving process. Typically, problem solving is explained as a five-step process:

1. Identify and define the problem
2. Identify alternative solutions to the problem
3. Select the one solution that best fits the problem
4. Install the selected solution
5. Evaluate the process of installation and the effectiveness of the solution

EPEC deals with steps four and five, installing and evaluating a selected problem solution. It is important that you understand the focus of this system at the outset.*

It is also important for you to understand that the developers of EPEC believe 1) that group interaction is the most appropriate way to communicate to you the systematic observation skills necessary for evaluating the process of educational change; 2) that EPEC's content will become part of you only if you practice your learning in both simulated and real settings. Thus, the EPEC modules will involve much group activity and practice.

You may experience some discomfort, frustration even, as you adapt to this style of learning. The time allotted for group activity may seem too much at first and too little later, as you proceed through the modules. This is to be expected. After all, you are now the consumers of an educational innovation. The instructor will listen to your concerns, but the developers have asked him not to make any changes in the sequencing and timing of EPEC module activities. The developers will listen to your concerns, too, as you record them on the EPEC Evaluation Form you receive at the end of each module.

STOP

* Participants who desire additional training in steps one, two and three of the problem solving process should arrange to take Research Utilizing Problem Solving (RUPS), an instructional system developed by the Northwest Regional Laboratory, Portland, Oregon.

APPENDIX C: EPEC OBJECTIVES

EPEC ACTIVITIES - MODULE 1

DESIGN: 1. Purpose
2. Rationale
3. Performance based objectives (level 1)

1. CALENDAR FOR EPEC MODULES

Purpose: Arrange meeting time to conduct group sessions

Rationale: Positive effect on modules if the participant can determine the time most suited to his schedule for the modules

Objective: EPEC participants will agree on and individually record on the proper page a schedule for the nine EPEC modules.

2. ORIENTATION TO EPEC

Purpose: Give the participant a perspective of the package

Rationale: Establish a common frame of reference to facilitate acquiring the objectives of the course

Objective: Participant will correctly identify during the Self-test
1. the purpose and objectives of EPEC
2. that EPEC's method of instruction focuses on group process techniques
3. that EPEC content emphasize the development of process evaluation skills

3. TAPE PRESENTATION AND ELEMENTS OF INNOVATION INVENTORY

Purpose: Present the participant with an organizational framework for processing information on an educational innovation

Rationale: Necessary competency of an evaluator

Objectives: 1. The EPEC participant will identify the following variables presented in the taped presentation by recording them on pp. I-11, I-12

a. the problem
b. the proposed solution
c. the arguments for and against the proposed solution
d. the advocate, the consumer, the innovation, and the organization in which the innovation will occur
e. the role of the evaluator in terms of these interacting variables

2. Participant experiences the dynamics of working in a team to further clarify the problem condition

4. GUIDELINES FOR WRITING A PROBLEM STATEMENT

Purpose: Give the participant guidelines for writing a problem statement

Rationale: Give the participant functional rules which he can immediately use to dilinate the problem conditions and then further develop to create the system best suited to his role as an evalutor

Objective: Participant will write a problem statement (I-18) according to the "Guidelines" given on p. I-17

5. DISCUSSION OF PROBLEM STATEMENTS

Purpose:
1. Discuss their problem statements
2. Build teamwork skills

Rationale:
1. Discussion gives the participant feedback on his interpretation of the material
2. Teamwork skills are:
   a. necessary for meeting the objectives of EPEC
   b. necessary competancy of a process evaluator

Objective: Each participant will modify his written problem statement as a result of input from the trio discussion outlined on p. I-19

6. FORCE FIELD ANALYSIS

Purpose: To give the participant a tool for analyzing the forces that have created a given state

Rationale: To obtain a better understanding of the problem state to facilitate reaching the goal state

Objectives: The participant will identify b correctly responding to the self-test (I-22):
a. the purpose of FFA;
b. the manner in which the "perceived situation" is represented in the FFA.
Participant will perform force field analysis on his problem statement using the "Guidelines" given on I-20, I-21, and I-24

7. DISCUSSION OF FORCE FIELD ANALYSIS

Purpose:
1. Discuss force field analysis
2. Build teamwork skills

Rationale: Feedback facilitates learning

Objective: Participant will modify his written problem statement as a result of input from the trio discussion outlined on p. I-25

8. ORIENTATION TO CONTRACT AND RULES FOR BRAINSTORMING

Contract

Purpose: State clearly for the participant what he is exp to do for his final project

Rationale: If the participant knows specifically what he is expected to do he will be better able to work independently on the project, as opposed to being dependent on the instructor for direction and control of the project

Brainstorming

Purpose: Give the participant a tool for generating a lot of ideas in a short amount of time

Rationale: An aid for diminishing the probablility of getting bogged down in the initial phase of a problem
1. Participants follow "Rules for Brainstorming"
2. Each participant identifies one contract topic during the brainstorming session for the final project in the course
3. Participants will identify the rules for brainstorming by scoring at least 50% on the self-test (I-30).

EPEC ACTIVITIES - MODULE 2

DESIGN:
1. Purpose
2. Rationale
3. Performance based objectives (level 1)

1. IN-BASKET EXERCISE # 1

Purpose:
1. To place the participant in a simulated decision-making role
2. Call upon the participant to make decisions based on the situations structured in EPEC

Rationale: The simulation will sensitize the participant to the role of decision maker bringing insight to his role as process evaluator

Objective: Participant will record a decision for each in-basket item on the Action Matrix

2. ACTION ANALYSIS PROFILE

Purpose: Organized form for listing possible actions taken at a decision-point

Rationale: To add to the participant's understanding of the decision role by alerting him to alternative decisions he may not have considered

Objective: Participant fills out the "Action Analysis Profile"

3. MEANS OF COMMUNICATION PROFILE

Purpose: Organized form for listing possible channels for communicating information following the decision

Rationale: To add to the participant's understanding of the decision role by alerting him to alternative channels of communication he may not have considered

Objective: Participant fills out the "Means of Communication Profile"

4. DISCUSSION OF DECISION ALTERNATIVES

Purpose: To have the participant experience the different solution strategies that can be generated by a group of people to the same problem situation

Rationale: Awareness of alternative solutions will increase the number of potential decisions in the thought process of each participant

Objective: Each participant will present his alternative solution strategies and will respond to those presented by his colleagues, according to "Guidelines" on p. II-15

5. EVALUATION DECISION MAKING AND INSTITUTIONALIZATION

The participant will read EDM and I (II-16 through II-20).

6. PROBLEM STATEMENT AND FORCE FIELD ANALYSIS REVISITED

Purpose: Reinforce the framework previously given for analyzing a problem situation

Rationale: Practice in using the framework will increase the skill level for use in the field

Objective:
1. Participant will revise his original problem statement using p. II-24 to record his revision.
2. Participant will revise his original force field analysis, using p. II-24 to record his revision.

7. FISHBOWL ACTIVITY

Purpose:
1. Organize group process for obtaining feedback
2. Identify some of the elements and relationships in the group discussion through the use of assigned roles
3. Work within a time limit
4. Build teamwork skills

Rationale:
1. Feedback functions to develop skills and increase productivity of both the individual and the group
2. Assigned roles aid a participant to focus on a specific task for the group
3. Working eithin a time limit assists the participant to focus on prime elements first
4. Teamwork skills are meant to increase productivity and facilitate the solution of problems

Objectives:
1. Each participant will implement the "Guidelines" (II-25, II-26) for the Fishbowl Activity.
2. Each participant will express satisfaction with the helpfulness of input received during the fishbowl

EPEC ACTIVITIES - MODULE 3

DESIGN:
1. Purpose
2. Rationale
3. Performance based objectives

1. IN - BASKET EXERCISE #2

Purpose:
1. Use the participant's understanding of the situation as a resource to create an on-the-spot problem incident
2. Process information from observing a problem incident into a problem statement
3. Reach consensus on adequacy of problem statements

Rationale: Use of simulation as a training technique to practice process evaluation skills

Objective:
1. The participant will formulate a problem statement on p.III-6 based on inputs from the in-basket exercise, p.III-5
2. The participant will reach consensus within his trio on technical adequacy of his problem statement--according to the instructions on p. III-7
3. The participant will reach consensus within his trio on the content adequacy of his problem statement--according to the instructions on p.III-7

2. REVISION OF THE CONTRACT

Purpose:
1. Obtain feedback on their contracts
2. Use new knowledge to revise contract

Rationale: Emphasize contract as their immediate vehicle to be used in applying knowledge as it is gained in EPEC to a practical situation.

Objectives:
1. The participants will discuss their contract with other members of the trio according to the guidelines on p. III-8
2. The participant will revise the contract (III-9, III-10) using input received during the trio discussion

---

OBJECTIVES FOR READINGS LISTED IN PACKAGE

EPEC ACTIVITIES - MODULE 4

DESIGN:
1. Purpose
2. Rationale
3. Performance based objectives

1. CHECKLIST OF INNOVATION CHARACTERISTICS

Purpose: Apply Brickell and Havelock's fremework for analyzing characteristics of an innovation to the CAI innovation

Rationale: To provide a useful technique for analyzing innovations which the participant will encounter in the field.

Objectives:
1. The participant will fill out the "Checklist of Innovation Characteristics" (p. IV-6) as per instructions
2. The participant will vocalize no objection to the scientific status of the EPEC instrument.

2. DISCUSSION OF INNOVATION CHARACTERISTICS

Purpose: Compare and discuss responses made on the checklist

Rationale: Participant needs feedback to clarify and strengthen his interpretation of the material

Objectives:
1. Each participant will give and receive feedback concerning responses to the checklist
2. The participants will reach consensus concerning those innovation characteristics whicy apply to CAI

3. IN-BASKET EXERCISE #3

Purpose:
1. To place the participant in a simulated decision making role
2. Call upon the participant to go through the decision making process
3. Analyze the decision making process in terms of the profiles given in EPEC

Rationale: Give insight to the decision making process

Objective:
1. The participant will make a decision on each in-basket item and record the decision on the "Action Matrix," p. IV-16

2. The participant will transpose his strategies to "Action Analysis Profile" and "Means of Communication Profile" pp. IV-17—IV-19
3. The participant will complete the "Situation Analysis Profile," p.IV-20.

4. DISCUSSION OF DECISION ALTERNATIVES

Purpose:
1. To analyze different solution strategies in terms of a common framework (3 analysis profiles)
2. To work within a time limit

Rationale:
1. Awareness of alternative solutions will inrease the number of potential decisions in the thought process of each participant
2. Common framework will facilitate assessing the adequacy of each strategy
3. Time limit for presentation should force participant to focus on key elements first
4. Working within a time limit should aid in keeping participants on task

Objectives: The participants will present and analyze in trios alternative strategies using the analysis profiles on p. IV-16—IV-20

5. PROBLEM STATEMENT AND FORCE FIELD ANALYSIS IN-BASKET # 3

Purpose:
1. Strengthen these skills in the participants repertoire of process evaluation skills
2. Stress these skills as the first steps to perform in analyzing new information

Rationale: Necessary for process evaluation

Objectives:
1. The participant will write a technically adequate problem statement (p. IV-22) on information obtained from in-basket #3 items
2. The participant will perform a force field analysis of problem statement on p. IV-22

6. FISHBOWL ACTIVITY

Purpose:
1. Organize group process for feedback
2. Help participants analyze team processes
3. Build teamwork skills

Rationale: Necessary competency of process evaluator

Objectives: The participants will follow the instructions for fishbowl activity (p. IV-23) using their problem statements and force field analyses as topics

7. REVISION OF CONTRACT

Purpose:
1. Revise contract based on new knwoledge gained from EPEC
2. Provide the participant with feedback on his contract from members in his group

Rationale: Further use of the contract as the vehichle for tying knowledge as it is gained to an actual innovation

Objectives:
1. The participant will revise the "Innovation" section of his contract
2. The participant will discuss the "Innovation" section of his contract with the members of his trio

---

OBJECTIVES FOR READINGS LISTED IN PACKAGE

EPEC ACTIVITIES - MODULE 5

DESIGN:
1. Purpose
2. Rationale
3. Performance based objectives

1. SUMMARY AND PERSPECTIVE I

Purpose:
1. Give the participant a check point to match the intended objectives of the package with what he has learned. If there is an apparent discrepancy he should be able to remedy it by going back over parts of the package. (The "EPEC Activities" can serve as an aid for the instructor in helping the participant locate the relevant activities.)
2. Bring the participants into the same frame of reference for proceeding with the rest of the package
3. Give the participants an overview of the model on which EPEC was designed

Rationale:
1. The check point reinforces those participants who have put the necessary work into the package. It allows those that haven't an opportunity to catch up and attain the objectives of the first three modules, which are prerequisites for the rest of EPEC
2. Common reference point or perspective is meant to help the participants attain the objectives for the following modules
3. Giving the participants the module on which EPEC is based provides them with one way to assess the value of the package and the rationale behind it.

Objectives:
1. The participant will read "Summary and PerspectiveI."
2. The participant will correctly identify in the Mastery Test a) that the second half of EPEC is concerned with the advocate and consumer rather than the organization and the innovation; b) that observing and reporting advocate-consumer-behaviors are a necessary skill of the process evaluator--that memorizing and reporting the behavioral categories is not sufficient.
3. The participant will write out the three purposes of EPEC on the Mastery Test.

2. INTRODUCTION TO ROLE PLAY AND ROLE PLAY ICEBREAKER

Purpose:
1. Explain purpose and technique of role playing
2. Give all the participants a chance to practice the technique in a simple exercise

Rationale: Necessary for the participants to have an understanding of the technique because they will be using it as a learning toll in following activities

Objectives:
1. The participant will read Introudction to Role Play, P. v-7
2. The participant will participate in the role play icebreaker followin the role play cards, (ppV-7b—d)

3. ROLE PLAY CALENDAR

Purpose: Solicit volunteers for the role playing activities

Rationale: Having the roles assigned in advance should help the activities start on schedule

Objectives: All participants will record the player's name for each character
Each participant will volunteer to play one role

4. ROLE PLAY # 1

Purpose:
1. To provide consumer and advocate the opportunity of simulating an actual confrontation
2. To focus on the consumer resistance to an innovation
3. To provide the remaining participants an experience in observing a consumer advocate confrontation

Rationale:
1. Simulation is an effective technique for actively involving participants in the learning experience
2. The simulation exercise will aid participants in their understanding of systems they will encounter in the field
3. Simulation is an effective training technique for building observation skills

Objectives: The participant portraying the consumer will enact consumer resistance characteristics based on the situation structured in EPEC
The participant portraying the advocate will try to sell the innovation based on information given about the consuemr and the resistant behaviors he is exhibiting
The participant serving as process evaluator will record observations of consumer behavior on practice sheet supplied

5. GUIDELINES FOR DISCUSSION OF FIRST ROLE PLAY

Purpose: Structure the discussion so that:
1. The participant playing the consumer gets feedback on the consumer behaviors he illustrated

2. Determine if there is a match between the behaviors he intended to illustrate and the behaviors the observers picked up
3. Determine if there is consensus among the observers on the behaviors recorded
4. Analyze any mismatches in the points above

Rationale:
1. Feedback on their observations assist them in building observation skills
2. The particular structure assists the participants to focus on the objectives of the exercise, while guarding against their spending too much time on insignificant points

Objectives:
1. The participant will follow the Guidelines for Discussion (p. V-15.)
2. The participant will indicate on the mastery test that EPEC instruments and checklists are intended only as learning tools--not as validated instruments

6. ATTITUDE INSTRUMENT AND TRIO ACTIVITY

Purpose: Give the participant an opportunity to assess his attitude toward role jplaying and discuss it with the other members of his trio

Rationale: If the participants more clearly understand their feelings towards role playing, it should assist them in analyzing future role playing activities

Objectives:
1. The participant will complete the "Summary Instrument" (p. V-16.)
2. The participant will discuss with his trio his responses to the summary instrument
3. The participant will verbally react to the summary instrument responses of the other members of his trio.

7. ROLE PLAY # 2

Purpose: Provide the participant an opportunity to study the role of process-evaluator through his active involvement in a role play exercise

Rationale: Effective training technique for building the skills necessary to a process evaluator

Objectives: The participant role playing the evaluator will use a strategy to assist the advocate in accordance with role play card (p.V-19) The participant role playing the advocate will seek help from the evaluator . The participant role playing the observer will complete an ovservation sheet (p.V-21.)

8. GUIDELINES FOR DISCUSSION OF ROLE PLAY # 2

Purpose: Structure the discussion so that:

1. Participant playing George High gets feedback on the strategy he used to service a decision maker
2. Analyze the strategy used by the role player and possible alternative strategies in terms of the concepts and skills taught in EPEC

Rationale: Assist the participant in his conceptualizing about the role of a process evaluator

Objectives: The participants will follow guidelines for discussion (p.V-22)

9. CONTRACT REVISION

Purpose:

1. Through discussion obtain feedback on present state of contract in terms of interview guidelines
2. Begin revision of contract

Rationale:

1. Feedback from other members will be useful in revisin contract
2. Begin revision while feedback from group is fresh in participant's mind

Objectives:

1. Participants discuss contracts
2. Participants begin revision on their contract to be presented in interview

DESIGN: 1. Purpose
2. Rationale
3. Performance based objectives

1. ROLE PLAY REVIEW

Purpose: Sensitize the participant to the complexity and interrelationships of variables he is attempting to delinate in his observations

Rationale: Noting and discussing discrepancies in his own observations of the same incident should assist him in building observation skills

Objectives:
1. The participant will list as the tape is played evidences of consumer resistance on the CRB Practice Sheet, p. VI-5.
2. The participant will discuss in trios their responses to the Practice sheet and the discrepancy between those responses and the responses made in Module 1, according to "Guidelines," p. VI-11.

2. ROLE PLAY # 3

Purpose:
1. Use the participants' understanding of consumer and advocate roles to simulate a confrontation to study consumer acceptance behaviors
2. Build observation skills

Rationale:
1. Awareness of consumer acceptance behaviors provides useful information for implementing an innovation
2. Observation skills are useful to the process evaluator

Objectives:
1. The participant playing the consumer role will portray consumer acceptance behavior consistant with the situation structured in EPEC
2. The participant playing the advocate role will try to gain more support for the innovation based on information given about the consumer and the behaviors he is exhibiting
3. The participant playing the process evaluator will record observations of consumer behavior on instrument supplied

3. GUIDELINES FOR DISCUSSION ROLE PLAY # 3

Purpose: Structure the discussion so that:
1. The participant playing the consumer gets feedback on the consumer behaviors he illustrated
2. Determine if there is a match between the behaviors he intended to illustrate and the behaviors the observer picked up
3. Determine if there is consensus among the observers on the behaviors recorded
4. Analyze any mismatches in the points above

Rationale:
1. Feedback on their observations assist them in building observation skills
2. Guidelines help them focus on the objectives of the activity

Objectives: The participants will follow the "Guidelines for Discussion" on p. VI-12

4. ROLE PLAY # 4

Purpose:
1. Demonstrate that it is likely that there will be both forces for and forces against the innovation operating in a consumer
2. Strengthen observation skills of participants

Rationale: Necessary to get a profile of the consumer that is as complete as possible for installing an educational innovation

Objectives:
1. The participant playing the consumer role will portray both consumer acceptance and resistance behaviors toward the innovation
2. The participant playing the advocate role will try to gain more support for the innovation based on knowledge of CAI and consumer behavior
3. The participant playing the process evaluator will record observations of consumer behavior on instrument supplied, (p.VI-16, VI-17)

5. GUIDELINES FOR DISCUSSION ROLE PLAY # 4

Purpose:
1. Feedback on the role play activity
2. Recognize that in focusing on both acceptant and resistant behavior, they have put together a force field analysis of a consumer's position which will need to undergo change before the goal state is reached
3. Discuss forces for and forces against during the role play in terms of their affect on installing the innovation

Rationale: A complete profile of the consumer in terms of a force field analysis has direct implications for formulating specific strategies to implement the innovation

Objectives: The participants will follow the "Guidelines for Discussion" on p. VI-18

6. SCHEDULE FOR CONTRACT AND INTERVIEW GUIDELINES

Purpose:
1. Schedule individual contract interview
2. Clarify the task by presenting guidelines for the interview

Rationale:
1. Formally obtaining agreement on an interview date puts the responsibility on the participant to be prepared
2. Guidelines will assist both the instructor and the participant to focus on performance relevant to the task

Objectives: The instructor and the participants will agree on interview dates and will record those dates p. VI-20

7. CONTRACT REVISION

Purpose:
1. Through discussion obtain feedback on present state of contract in terms of interview guidelines
2. Begin revision of contract

Rationale:
1. Feedback from other members will be useful in revising contract
2. Begin revision while feedback from group is fresh in participant's mind

Objectives:
1. The participants will discuss contracts in trios using VI-19 as a discussion guide
2. The participant will begin revision of his contract to be presented in the interview with the EPEC instructor

EPEC ACTIVITIES - MODULE 7

DESIGN:
1. Purpose
2. Rationale
3. Performance based objectives

1. ROLE PLAY # 5

Purpose: Use simulation to study the advocate strategies for bringing about adoption of an innovation as proposed in EPEC

Rationale: Training technique to assist the participants in assimulating the strategies presented in EPEC

Objectives:
1. The participant role players enact the role play characters described on the role play card.
2. The participant serving as process evaluators will record their observations of the role play using the "Advocate Strategies: Practice Sheet" (p. VII-7)

2. GUIDELINES FOR DISCUSSION - ROLE PLAY # 5

Purpose: Analyze role play focusing on advocate strategies

Rationale: Lead to a clear understanding of the advocate strategies proposed in EPEC

Objectives: The participant leader and discussants will follow the "Guidelines for Discussion" on p. VII-8

3. ROLE PLAY # 6

Purpose: Use simulation to integrate the concepts for advocate strategies with consumer resistant behaviors

Rationale: Method for analyzing interaction between consumer and advocate during the change process

Objectives:
1. The participant role players enact characters based on information given them
2. The participants serving as process evaluators record their observations of the role play using three practice sheets (p. VII-12, 13, 14)

4. GUIDELINES FOR DISCUSSION - ROLE PLAY # 6

Purpose: Analyze the interaction between advocate strategies and consu consumer resistant behaviors in terms of the constructs given in EPEC

Rationale: Necessary for the participants to understand the relationship between the parts of the model presented in EPEC, in order for it to be functional for them in the field

Objectives: The participant leader and discussants will follow the "Guidelines for Discussion" on p. VII-15

5. GUIDELINES FOR DISCUSSION OF ADVOCATE STRATEGIES

Purpose: Participants receive feedback from each other on the use of advocate strategies in their contract

Rationale: Enable the participants to make the most effective use of advocate strategies for installing their innovations

Objectives: Trio participants will discuss the advocate strategies which they listed in their contracts following discussions guidelines on p. VII-16.

DESIGN: 1. Purpose
2. Rationale
3. Objectives

1. SUMMARY AND PERSPECTIVE II

Purpose:
1. Review the instructional model on which EPEC was developed
2. Assist the participants to synthesize the skills and concepts for change presented in EPEC
3. Increase the probability that the model for change presented in EPEC will guide the installation of the change designed in their contract

Rationale: Increase the probability for success in operationalizing their contracts

Objectives: The participant will read "Summary and Perspective II" (p.VIII-3)

2. ELEMENTS OF INNOVATION INVENTORY

Purpose: Reinforce the framework given in EPEC for analyzing an innovation

Rationale: Practice in using the framework will increase the skill level for use in the field

Objectives: The participant will complete the inventory (pp.VIII-5 through VIII-7) based on the simulated CAI innovation

3. IN-BASKET EXERCISE # 4 AND DISCUSSION OF DECISION ALTERNATIVES

Purpose: Give the participants a mechanism for determining the effect of EPEC experiences upon their decision making style

Rationale: Awareness of changes in the participants' decision making style due to EPEC is likely to reinforce skills they have learned in EPEC

Objectives:
1. The participants will complete the Action Matrix, Action Analysis Profile, and Means of Communication Profile, and discrepancy sheet, according to the instructions provided on pp. VIII-16 through VIII-19.

2. The participants will discuss the effect EPEC has had on their decision making style, according to instructions on p. VIII-20.

4. REVISION OF THE CONTRACT

Purpose:
1. Get final feedback on their contract from group members before implementing their innovation
2. Analyze the contract for possible problem areas in implementation
3. Give mutual reinforcement for the work done and support for the job ahead

Rationale: Increase the probability for success in installing the innovation

Objectives:
1. The participant will discuss his contract and write down the parts of his contract which he perceives will be the most difficult to perform
2. The participant will obtain the instructor's consent for any revisions he wishes to make in his contract at this time

DESIGN: 1. Purpose
2. Rationale
3. Objectives

1. ORAL REPORT AND ASSESSMENT

Purpose: 1. Have the participant present to the group the outcome of his contract
2. Have participants perform the external evaluation of individual projects, based on the information given them in the presentation

Rationale: Mechanism through which the participant demonstrates accountability for his contract

Objectives: 1. The participant will present an oral report on the outcome of his contract
2. The participant will evaluate the oral reports of his peers, using the Oral Report Assessment Form, p. IX-3.
3. The participant will use at least five EPEC idiosyncratic terms in his oral report.

Entry Behavior: Participant will have completed VIII-27 through VIII-29 before Module IX begins

2. GUIDELINES FOR FINAL DISCUSSION

Purpose: 1. Enable the participant to discuss the evaluation he received on his oral report with the evaluators
2. Provide the participants with the opportunity to further explore any of the outcomes of the contracts
3. Share perceptionsof EPEC

Rationale: 1. Clarification of the evaluation report increases its relevance to the participant
2. Verbalizing their perceptions of EPEC should assist them in evaluating their "EPEC" experience

Objectives: The participant will follow the "Guidelines for Final Discussion" on p. IX-5

APPENDIX D: INSTRUMENTS

BACKGROUND INFORMATION

Name _____ Age _____ Sex _____

General Occupation _____ Years Professional Ex. _____

Current Position _____

Institution and Location _____

Telephone: Business _____ Home _____

What is your educational background? (Check all that apply)

_____ High School graduate

_____ Some College

_____ College Graduate (major) _____ (degree) _____

_____ Masters Degree (major) _____ (degree) _____

_____ Ph. D. Degree (major) _____ (degree) _____

_____ Current Student (major) _____ (degree) _____

(area of specialization) _____

How many quarter hours, if any, have you completed beyond your last degree? _____

Have you previously participated in a workshop or course involving any of the following? (check all that apply)

_____ Role Playing

_____ Brainstorming

_____ Sensitivity Training

_____ Institutional Change Theory

TEST #

Why are you taking this course and what do you expect to gain?

DIRECTIONS FOR ADMINISTERING THE PROBE QUESTIONNAIRE

During a training experience participants may react differently to each particular LA. Hence, if reactions are assessed at the end of a module of LA's much information tends to be lost due to forgetting or confounding. However, if participants must complete instruments after every activity too much time may be devoted to evaluation, participants may react negatively to the evaluation or the evaluation may interact with training producing different outcomes than would be obtained by training alone. The PROBE questionnaire is designed to tap participants' reactions to each activity on the spot with a minimum of interference and time. It is indeed a quick "probe" of participant reactions to training in terms of a very few but critical dimensions.

The PROBE questionnaires, one for each module, are attached to the inside back cover of the participant's workbook by means of a spring clip. The PROBE General Instructions is included in the workbook as the last paper to serve as a ready reference as participants need to check the meaning of the PROBE questionnaire.

Before training begins, participants are instructed to read the PROBE General Instructions and ask questions of the instructor. Then, at times designated by the instructor, participants are asked to complete PROBE for a specified activity. The instructor's agenda contains the cue, "PROBE," in stamped red letters, whenever a probe is to be made during training. Upon completing a step preceding the cue, the instructor gives participants the following directions:

"Please turn to the back of your workbook and complete PROBE for learning activity 1*." (*The instructor will, of course, give the number of the activity just completed. That number will follow the cue as follows:

PROBE: 1)

Some delay may be expected the first one or two times participants complete PROBE, but generally it is expected that training may proceed within a minute of the directions quoted above.

After completing PROBE on the last activity of a module, the instructor directs participants to complete the LOG on the questionnaire by saying, "Now complete the PROBE LOG by going back over your ratings and explaining any 'questionable' or 'not satisfactory' ratings. Please be as specific as possible about what you think should be changed." Five minutes will be allowed for this step. The PROBE instrument is then collected, leaving the PROBES for the coming modules in place.

Data from PROBE may be key punched directly from the questionnaire. Card columns are indicated next to each response box. The number inside the box is punched into the column indicated. The results of the data analysis are displayed in their final form by a custom-make computer program.

PROBE General Instructions

In order to help us evaluate EPEC we will need to know how you feel about the EPEC learning activities. You will therefore be asked to respond to the PROBE questionnaire at designated times during the course. Please scan your copy of the questionnaire, then read the following explanations. They will help you understand the kind of information we need.

THE MEANING OF THE SCALE

1 = superior: you believe this activity is outstanding in regard to the item rated
2 = satisfactory: the activity does not require revision
3 = questionable: the activity is almost adequate; you would recommend revision only if resources permit
4 = not satisfactory: you strongly recommend that the activity be revised

THE ITEMS

The meaning of the PROBE items will be illustrated by examining a hypothetical role playing activity. Assume part of your group had taken assigned roles in a meeting between an administrator who is advocating a new program and several teachers. The teachers were instructed to either resist or accept the change he advocated. The rest of the class observed in order to gain the skill of identifying consumer resistance and acceptance behaviors. Assume the activity is now over and you are about to complete the PROBE questionnaire:

Please rate the extent to which:

(a) The concepts/skills were relevant: In our hypothetical example the content included the skill of identifying consumer resistance and acceptance behaviors. Do you regard this skill as important? If so, you would check "satisfactory" or "superior"; if not, you would check "not satisfactory" or "questionable" suggesting the activity be revised.

In general, this item asks, "Is the content of this activity, including ideas, concepts, principles, skills, etc., of interest or important to you or would you agree they are important to the evaluation of change?

(b) The material was presented so you can understand it: Was the content presented in a clear and understandable manner-free of jargon and complexities? Remember, content includes ideas, concepts, principles, skills, etc. In our example the content was the skill of identifying consumer resistance and acceptance behaviors. Was this skill presented clearly--or would recommend revision?

(c) You were provided with helpful feedback on your progress: Did you receive adequate useful information concerning your mastery of the concepts/skills? Were there explicit procedures to remedy weaknesses? Did the feedback procedures facilitate learning or were they disruptive?

(d) Conditions for success were sufficiently specified: Were all the conditions necessary for success sufficiently spelled out? Perhaps you think there should have been a reading assignment on observation techniques, or that students need a warm up before they can jump into a role playing activity. If so, you would check "questionable" or "not satisfactory." The idea is to ask yourself, "what should have been done ahead of time to make this activity more successful?"

Rate also:

(e) The way the material was presented, the methods/media used: For this activity you would rate role playing as a way of presenting consumer resistance and acceptance behaviors. Was it appropriate? Did it work? Were there too many difficulties? Were instructions clear? Did you find it enjoyable and stimulating.

(f) The quality of the instructional materials: Were they ledgible, appealing, convenient, durable, etc.?

(g) The quality of the learning environment: Were the facilities appropriate and comfortable? Was the temperature satisfactory? Were you free of noise and distractions? Did you have enough time? Was the environment conducive to productive learning?

THE LOG

Explain your negative ratings: The scales above do not tell us what to change when you indicate an activity is not satisfactory or questionable. Use this space to explain any questionable or not satisfactory ratings you give. The more specific you can be about what to change, the more useful your ratings will be. You will not complete the log after each activity but only after each module (a set of activities.) You may, however, wish to make a cryptic note to yourself at the time you rate an activity to help you remember what you meant by a rating when you complete the log later.

GENERAL INSTRUCTIONS FOR COMPLETING THE SCALE

Give your first impression: Respond quickly to the rating scale; 30-40 seconds should be sufficient. Time will be saved and your first impressions are generally more valid than prolonged reflection.

Please respond to every item: If you have difficulty choosing between two alternatives just check the one that more closely reflects your feelings. If you missed part of an activity or don't feel you can rate the activity for some reason, then don't respond to the questionnaire at all.

Feel free to share your reactions with the instructor during breaks or after class. Remember, EPEC is being evaluated--not you. Your grade will not be influenced by your ratings.

Thank you, what we need money can't buy—we need you!

A-7

PROBE

Directions: Place your ratings in the boxes under the activity just completed.

SCALE: 1 = superior
2 = satisfactory
3 = questionable
4 = not satisfactory

ACTIVITY

Please rate the extent to which:

	#1	#2	#3	#4	#5	#6	#7	#8	#9	#10
a. the concepts/skills were relevant	☐11	☐18	☐25	☐32	☐39	☐46	☐53	☐60	☐67	☐74
b. the material was presented so you can understand it.....	☐12	☐19	☐26	☐33	☐40	☐47	☐54	☐61	☐68	☐75
c. you were provided with helpful feedback on your progress....	☐13	☐20	☐27	☐34	☐41	☐48	☐55	☐62	☐69	☐76
d. the time alloted was appropriate (rate 2 or 4 only)	☐14	☐21	☐28	☐35	☐42	☐49	☐56	☐63	☐70	☐77
Rate also:										
e. the way the material was presented, the methods/media used	☐15	☐22	☐29	☐36	☐43	☐50	☐57	☐64	☐71	☐78
f. the quality of the instructional materials	☐16	☐23	☐30	☐37	☐44	☐51	☐58	☐65	☐72	☐79
g. the quality of the learning environment	☐17	☐24	☐31	☐38	☐45	☐52	☐59	☐66	☐73	☐80

LOG: (Explain your negative ratings. Begin your statements with the number of the box you are writing about.)

INSTRUCTOR PROBE

Directions: Place your ratings in the boxes under the activity just completed.

SCALE: 1 = superior
2 = satisfactory
3 = questionable
4 = not satisfactory

ACTIVITY

Please rate the extent to which:

a. the concepts/skills were adequately covered.....

b. participants had the necessary entry behaviors...

Rate also:

c. the way the material was presented, the methods/ media used.....

d. the appropriateness/adequacy of the materials....

e. the appropriateness of the time allotment.....

f. the quality of the learning environment.....

LOG: (Explain your negative ratings)

Please rate the extent to which the module's activities formed a well organized whole as opposed to being disjointed or isolated with no sense of order or purpose:

_____Superior

_____Satisfactory

_____Questionable

_____Not Satisfactory

Explain:

Do you think the order in which the learning activities were presented was satisfactory?

_____Superior

_____Satisfactory

_____Questionable

_____Not Satisfactory

Explain:

General Comments:

DEFINITION AND EXPLANATION OF TERMS

Agenda Item

This is the step on the Instructor's Agenda that is currently being observed. Merely write the agenda step number you are observing in the column.

Time Allotted

Write the number of minutes allotted to this step (as it appears in the Instructor's Agenda).

Time Record

Record the time (from a clock or wristwatch) that the group begins each agenda item. Then the time actually consumed for each step can be determined at a later time.

Substantive Input Provided?

Respond "yes" if the instructor provides any substantive input during the observed step. Substantive input refers to any concepts, ideas, principles, skills and the like that relate directly to the material covered. If the instructor provides none of this kind of input but only engages in managerial-related behaviors then respond "no." If the instructor did nothing more than read, repeat and clarify instructions, for example, you would respond "no." Whenever you respond "yes" then note under explanations the nature of the input provided.

Instructor Control Provided?

Control behaviors are instructor-initiated activities that influence the group of participants so as to help them accomplish their task. The instructor might offer encouraging or reinforcing remarks when participants are on the right track. He might paraphrase to facilitate understanding. He might redirect participants when they drift off task. Respond "yes" whenever the instructor provides any control behaviors during the observed step that are not part of the instructor's instructions. Then under explanations note the control behavior and, if possible, the situation that the instructor is responding to. If no control behaviors occur, respond "no."

Objectives Met?

Refer to the objectives of each learning activity that you observe. Then make a judgment about whether they were satisfactorily met. If you think so, respond "yes." If one or more are not met, write "no" and explain the problem under comments.

Deviations?

Respond "no" if there are no departures from the instructions provided and "yes" if there are. Generally, other columns on the observation schedule are intended to catch most departures, such as time, and instructor input, so use this column as a catch-all for deviations or discrepencies that don't fit elsewhere.

Instructions Clear?

If it appears that participants are able to understand and follow correctly the instructor's instructions then write "yes." If not, write "no" and note the

problem under explanations.

Explanations and General Comments

This column has two functions: to explain and describe trouble spots identified in the previous columns, and to record any other noteworthy observations. In particular, any environmental conditions that hindered productiveness of the group are to be noted here. Include also such things as wrong page numbers, missing materials and reversed pages.

GENERAL INSTRUCTIONS

Don't fall behind

It is important to remain current in your observations so you don't miss and have to remember back to past steps. If you can't get all the information down, make quick notes of the major points and fill them in later. (leave space!) You can use the tape recorder to catch inportant exchanges and comments--talk into it yourself if you need to get something down fast.

Observe the trios when they separate

You may not be able to observe every trio but do observe some--take the tape recorder with you if possible. Continue to note trio departures from time lines and from instructions; note any difficulties they may have.

Observe departure conversations

Some of the most important remarks from participants and instructors will come at the conclusion of the module. Try to tape record and make comment on these.

In general, your job is to capture the essence of what is happening in EPEC. Your comments on problems with the Observational Schedule and your task as well as advice on ways to improve them will be appreciated.

D-14

OBSERVATION SCHEDULE

Observer's name _____

Instructional System _____

Date _____ Module# _____

Agenda Item	Time Allotted	Time Record	Substantive Content Provided?	Instruction Control Provided?	Objectives met?	Deviations?	Instructions Clear?	Explanation and General Comments

DIRECTIONS

This instrument is designed to determine the types of skills a student is exposed to during EPEC.

Read the EPEC scale carefully, then rate each item according to the contribution you feel this course made to your competence in that area by circling the appropriate number.

EPEC

1. Not Applicable--EPEC was not designed to affect student competence in this area.

2. No Effect--Although EPEC was presumably designed to affect student competence in this area, it failed to do so.

3. Minimal Effect--EPEC had some effect on student competence in this area.

4. Large Effect--EPEC was highly beneficial in promoting student competence in this area.

5. Complete Mastery--As a result of EPEC, students are fully qualified in this area.

1 2 3 4 5 1. Can interact effectively with a decision maker in an evaluation setting.

1 2 3 4 5 2. Can function as a change agent in the educational change process.

1 2 3 4 5 3. Can identify alternative ways to handle a decision situation, such as "defer to higher authority" or "postpone decision"

1 2 3 4 5 4. Can use some effective group interaction tools such as brainstorming, fishbowl and role playing.

1 2 3 4 5 5. Have developed group observation skills.

1 2 3 4 5 6. Can describe the kinds of transaction which occur between advocates and consumers.

1 2 3 4 5 7. Can develop, select, and use process evaluation instruments.

1 2 3 4 5 8. Can analyze and report major results of data selected.

1 2 3 4 5 9. Can utilize concepts and skills of giving and receiving feedback.

1 2 3 4 5 10. Can determine the effectiveness of an implemented innovation.

1 2 3 4 5 11. Can formulate a problem statement.

1 2 3 4 5 12. Can perform a force field analysis.

1 2 3 4 5 13. Can identify strategies or action alternatives from a force field analysis.

1 2 3 4 5 14. Can install a specific educational change .

1 2 3 4 5 15. Can explain the difference between installing an innovation and evaluating the installation of an innovation.

1 2 3 4 5 16. Can list and describe advocate strategies.

1 2 3 4 5 17. Can list and describe organizational characteristics.

1 2 3 4 5 18. Can describe the role of the process evaluator.

1 2 3 4 5 19. Can evaluate the process of educational change.

1 2 3 4 5 20. Can identify the kinds of decisions an evaluator services.

1 2 3 4 5 21. Can describe an innovation in terms of innovation characteristics.

1 2 3 4 5 22. Can list and describe the levels of acceptance of an innovation.

1 2 3 4 5 23. Can list and describe the stages of adoption of an innovation.

1 2 3 4 5 24. Can assess the level of acceptance of an innovation.

1 2 3 4 5 25. Can assess the stage of adoption of an innovation.

1 2 3 4 5 26. Can list and describe consumer acceptance and resistance behaviors.

1 2 3 4 5 27. Can recognize consumer acceptance and resistance behaviors in human interaction.

1 2 3 4 5 28. Can assess the extent of consumer acceptance and resistance toward an innovation.

You have accepted a position as process evaluator for an elementary principal who is implementing a new reading program! Describe what you would do in this situation.

APPENDIX E: ANALYSIS FORMS

TABLE 1

MODULE SUMMARY

Data Indicating Revision is Necessary

Element	Activity									
	#1	#2	#3	#4	#5	#6	#7	#8	#9	Module
Content										
Objectives										
Environmental Conditions										
Media										
Materials										
Feedback										
Time										
Placement										

Y

Data indicate that criteria are met: Revision not needed.
Insufficient Data
A Not Applicable

rticipant Data from PROBE and the Module Questionnaire

TABLE A-1
TABLE A-2
TABLE A-3
DISPLAY A-1

structor Data from the Instructor PROBE and Log.

TABLE B-1
DISPLAY B-1

bserver Data from the Observation Schedule

TABLE C-1
TABLE C-2
DISPLAY C-1

TABLE A-1

PROBE Summary: Module

Proportion Checking Satisfactory or Superior

Activity	Variable							
	Relevant	Understand	Feedback	Environmental Conditions	Media & Methods	Materials	Time	$\bar{X}$
1								
2								
3								
4								
5								
6								
7								
8								
9								
10								
$\bar{X}$								

Participant Open-Ended Remarks: Module

:tivity #	Variable	Participant	Remarks

TEST #

TABLE A-2

PROBE Results: Module Activity

	Variable						
	Relevant	Understand	Feedback	Environmental Conditions	Methods	Materials	Time
Frequency Distribution							
Superior							
Satisfactory							
Questionable							
Not Satisfactory							
Culumative Percentage Distribution							
Superior							
Satisfactory							
Questionable							
Not Satisfactory							

TABLE A-3

Module Questionnaire Frequency Distribution Module

Rating	Variable	
	Continuity	Placement
Superior		
Satisfactory		
Questionable		
Not Satisfactory		

TABLE A - 4

E-6

Performance Test Results: Module 1

Item	f Responding	f Correct	% Correct

TEST #

TABLE A - 5

Performance Test Frequency Distribution

Score	f	%	cum %	$\bar{X}$	S.D.

TEST #

TABLE B-1

Instructor PROBE Ratings: Module

	Variable					
Activity	Content	Environmental Conditions	Media & Methods	Materials	Time	Entry Behaviors
1						
2						
3						
4						
5						
6						
7						
8						
9						
10						

DISPLAY B-1

Instructor PROBE LOG Summary: Module

Activity	Variable	Remarks

TABLE C-1

Observation Data: Module

Activity		Variables Needing Attention					
#	Description	Time Match	Substantive Content Provided?	Instructor Control Provided?	Objectives Met?	Instructions Clear?	Other Deviations

TABLE

Attendance Record

Module	Date	Participant #					
		01	02	03	04	05	06
1							
2							
3							
4							
5							
6							
7							
8							

DISPLAY

Test Site Information

Site #	N	Location	Instructor	Starting Date	Test Stage

DISPLAY

List of Participants

Trio #1	I.D.#	Business Phone	Home Phone
1.			
2.			
3.			
Trio#2			
1.			
2.			
3.			

APPENDIX F: DATA ANALYSIS INSTRUCTIONS

DATA ANALYSIS INSTRUCTIONS

STAGE 1: ANALYSIS FORMS

Complete the data analysis forms in the following sequence:

Table A-2: Provide a frequency distribution and cumulative percentage distribution of the participant PROBE ratings for each activity.

Table A-1: For each variable of each activity provide the cumulative percentage at the satisfactory level as obtained from Table A-2.

Display A-1: List the participants' comments from PROBE and the Module Questionnaire grouped according to variable and activity.

Table A-3: Provide a frequency distribution of participant responses to the Module Questionnaire.

Table A-5: List each performance test item along with the number of participants responding to the item, the number of correct responses and the per cent of correct responses.

Table A-4: Provide a frequency distribution of total scores.

Table B-1: Provide a frequency distribution of Instructor PROBE ratings for each activity.

Display B-1: List the comments from the Instructor PROBE grouped by element and activity.

Table C-2: Provide the information called for by the column headings (enter yes or no) for those agenda items needing attention. Under the heading "time match" indicate the time discrepancies for agenda items, for example, +10 minutes or -15 minutes.

Table C-1: Condense the information in Table C-2 so as to summarize the results by activity, as opposed to an agenda item summary.

Display C-2: List the observer's comments grouped by agenda item.

STAGE 2: EXAMINING THE ELEMENTS OF ACTIVITIES

A. Identify the criterion variable of interest to the particular element under examination. (Table D provides a list of the criterion variables pertaining to each of a learning activity's elements.)

F-2

TABLE D

Information Sources for the Data Analysis

Criterion Variable	Decision Rule	Analysis Form	
1. Substantive content of the LA:			
a. must not be provided by the instructional manager	f provided content=0	Table C-1	Display C-1
b. should be perceived by participants as relevant	positive response≥85%	Table A-1	Display A-1
2. Purpose of the LA: (none)			
3. Performance-based Objectives of the LA:			
a. must be attained	Per Cent (x≥90%)≥90%	Table A-4	
4. Media and Methods of the LA:			
a. must be perceived as satisfactory by			
1. participants,	affirmative response≥85%	Table A-1	Display A-1
2. the instructional manager.	affirmative response	Table B-1	Display B-1
b. must be implemented as specified.	f deviations=0	Table C-2	Display C-1
5. Materials of the LA:			
a. must be perceived by participants as acceptable*	affirmative response≥85%	Table A-1	Display A-1
b. must be perceived by the instructional manager as appropriate	affirmative response	Table B-1	Display B-1
6. Learning Environment Specifications of the LA:			
a. must be perceived as implemented by the observer	f departures=0		Display C-1
b. must be perceived as necessary and sufficient by			
1. the observer,	f reported problems=0		Display C-1
2. the instructor, and	affirmative response	Table B-1	Display B-1
3. the participants.	affirmative response≥85%	Table A-1	Display A-1
7. Entry Behaviors of the LA:			
a. must be perceived as necessary, sufficient and met			
1. the observer, and	f reported problem=0	Table C-2	Display C-1
2. the instructional manager	affirmative response	Table B-1	Display B-1

*See test criteria for a breakdown of this category.

TABLE D (cont.)

Criterion Variable	Decision Rule	Analysis Form	
8. Time allotment of the LA:			
a. must correspond with the actual time consumed	f discrepancies=0	Table C-2	Display C-1
b. must be perceived as appropriate by			
1. participants, and	affirmative response≥85%	Table A-1	Display A-1
2. instructional manager	affirmative response	Table B-1	Display B-1
9. Feedback Procedures of the LA:			
a. must be perceived by participants as adequate and not disruptive	affirmative response 85%	Table A-1	Display A-1
10. Placement of the LA:			
a. must be perceived by participants as appropriate in terms of sequence and continuity	affirmative response 85%	Table A-3	Display A-1

Example: For element 1, "Substantive Content of the LA" the first criterion variable is "must not be provided by the instructional manager."

B. Locate the relevant data from the analysis forms for the criterion variable being examined. (Table D lists the analysis forms containing the data appropriate for each criterion variable)

Example: For the criterion variable selected Table D refers the evaluator to Table C-1 and Display C-1. Here he will find data pertaining to the first criterion variable.

C. Review the data available and determine whether the criterion is met according to the decision rule specified in Table D.

Example: An examination of Table C-1 will reveal either a "yes" or a "no" under the heading "substantive content provided?"

D. If the criterion was met, proceed to the next criterion variable of the element being examined and repeat steps A through C. When examination of a particular element is completed and all criteria were met, then enter a "+" in the appropriate cell of Table 1, the module summary.

Example: Under the heading "substantive content provided?" a "no" has been entered. The decision rule in Table D states that the criterion is met if the frequency of substantive-content-provided-during-the-activity is zero. Since no substantive content was provided the criterion is met. Therefore, the evaluator proceeds to the second criterion variable under the element "substantive content of the LA," which is, "should be perceived by participants as relevant," and repeats steps A, B and C. If this criterion variable is also satisfactory the evaluator enters a "+" in Table 1 of the module report under the column of the activity being considered and in the row headed "content." He then proceeds to the next element "Performance-based objectives of the LA," and repeats steps A through D again.

E. If during step C, the criterion was not met, then instead of a "+", enter the appropriate code number in Table 1. The code number represents the analysis form that contains the information suggesting revision is needed. The key to the code is displayed at the bottom of Table 1. Providing this coded entry in Table 1 allows someone reading the module report to locate the source of any data suggesting a need to revise an activity.

Example: If, under the column of Table C-1 headed "substantive content provided?" the evaluator found a "yes" for the activity under consideration, then he would enter in Table 1 a "7", the code for Table C-1.

F. If it is not possible to determine from the data obtained whether the criterion is met then a "-" should be entered in the appropriate cell of Table 1 indicating insufficient data.

There may also be times when a particular element does not apply to the activity under consideration. In this event, an N/A should be entered in Table 1.

G. Whenever a criterion is not met and there is evidence suggesting revision, that evidence should be discussed in the text of the module report. This discussion should summarize and evaluate the strength of the evidence for revision. An effort should be made to relate other obtained information to the case being considered in order to better understand the nature of the problem. The evaluator might listen to the tape recording of the activity, for example, to get a better idea as to what is not satisfactory. He should also discuss any open-ended remarks that relate to the problem. If there are arguments against revision of the activity, these also should be considered. An effort should be made to explain what is causing the problem whenever this has implications for how the activity might be improved.

H. The final step is to recommend how the activity might be revised to strengthen it or alleviate the problem, or even to recommend that the evidence is not sufficient to warrant revision.

MODULE REPORT CONTENTS

The report should begin with the following opening paragraphs:

> The results of this module are summarized in Table 1. Each activity is considered in terms of its component characteristics or elements, that is, the activities content, objectives, materials, etc. A + indicates that no evidence was obtained to suggest the element of the specified activity needed to be revised. One or more digits indicates that there is evidence suggesting the activity be revised. The digits are codes that refer the interested reader to data summaries and comments in Appendix A that constitute the evidence for revision. The code, and instructions for its use, also appear in Appendix A.
>
> In the following discussion we will dispense with consideration of activities judged OK and consider the evidence for and against revision of the remaining activities.

Table 1 follows the above paragraphs. The next section of the report is the Discussion and Recommendations. This section should be organized by learning activity, and within learning activities it should be organized by element.

The last section of the report is a summary of the recommendations. The analysis forms are included as Appendix A and the raw data is included as Appendix B.

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted] R & D Report No. 1977-3

University of Minnesota
College of Education
Education Career Development Office
Minneapolis, Minnesota

April, 1977

(Presented at the 1977 AERA Annual Meeting: New York)

Abstract

This research investigated stochastic approximation procedures of the Robbins-Monro type. Following a brief introduction to sequential experimentation, attention is focused on formal methods for selecting successive values of a single independent variable. Empirical results obtained through computer simulation are used to compare several formal stochastic approximation techniques and "stopping rules". Marked differences were found between the five approximation procedures studied. One procedure using a "finite memory" had utility. Two procedures suggested in the literature were impractical under various conditions.

Introduction

Characteristically, an experiment involves a collection of treatment conditions (i.e. treatment levels or treatment combinations), a collection of experimental units, and an explicit plan for assigning treatment conditions to units. For purposes of exposition, we can divide experiments into those in which time plays an important role and those in which it does not. Time may enter into the experimental plan in several ways. For example, 1) at some point during a sequence of repeated measurements of the experimental unit a treatment condition may be introduced, as in trend analysis, 2) the experimental material may be successively exposed to several pre-specified treatment conditions and measured after each, as when assessment of order or residual effects is of interest, 3) treatment conditions may be administered to experimental units over time in such a way that previous treatment conditions and responses to them are used in determining the treatment conditions which follow. Note that in examples two and three treatment conditions are administered over time. But in the second example the exact treatment conditions are determined a priori, while in example three, they are determined during the experiment as a function of accumulating data. For convenience, we label the three examples as instances of repeated measurement, serial, and sequential designs, respectively.

The present research is concerned with sequential experimentation. Experimental designs which are sequential in nature require that the experimenter consider both how the ensuing treatment conditions will be changed or adjusted and how the process will be discontinued,

i.e. a "stopping rule". Sequential experiments can be differentiated from one another by considering whether a formal or informal procedure is used when adjusting treatment conditions, whether the stopping rule is formal or informal, whether or not more than one factor is used (i.e. multifactor experiment employing several different treatments), whether the independent variable or dependent variable is continuous or discrete, and by considering the purpose of the procedures (e.g. locating maxima). (For a general review and bibliography of recent work on experimental design, including the topics dealt with here, see Herzberg and Cox (1969). For a current review of the design of sequential experiments, see Chernoff (1975). Wetherill (1975) provides a useful introduction to the subject of our paper).)

Examples of applications of sequential designs are not plentiful in the educational research literature. Meyer (1963) presents an application of response surface methodology. This methodology is seen as sequential in nature by Chernoff and by Wetherill. Response surface designs are factorial in nature, employing several quantitative independent variables. The dependent variable is often assumed continuous and a polynomial function of the independent variables. Purposes of these designs include locating maxima or estimating parameters of the polynomial. Decision rules which specify the "design points" to use in the next stage and when to stop the process tend to be informal.

In contrast to response surface methodology are stochastic approximation techniques in which a single continuous independent variable is investigated and where values of that independent variable are determined formally as a function both of the preceding values and

the responses that were obtained when they were administered. A technique due to Robbins and Monro (1951) is an example. Its purpose is to find that value of the independent variable, say θ, such that the expected value of the dependent variable given θ is equal to some predetermined constant.

Our research investigates two examples of the Robbins-Monro process and three variant procedures which were motivated by it. Much of the previous research in this area has been focused on asymptotic properties. Chernoff (1975) gives a brief and readable review of this work. Of particular interest here is a paper by Hodges and Lehmann (1956) because it suggests assuming a linear relationship between the independent and dependent variables and also assumes that the slope parameter is known. While these two conditions would seldom be met in practice, their theoretical and numerical results provide a basis of comparison for empirical findings.

The Robbins-Monro procedure has been modified by some researchers so that two values of the independent variable are employed at each step (e.g. see Venter (1967)). This procedure has certain advantages, but only the case in which a single value of the independent variable is used at each step is studied here.

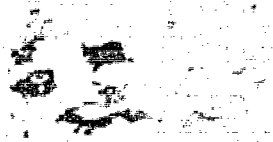
Presently, we know of no application of Robbins-Monro procedures in an educational experiment. However, the technique has been applied to a measurement problem by Lord (1971a, 1971b). Those two papers dealt with quantal responses, a subject not dealt with here. (For this reason and because we did not want to define the values of the independent variable a priori we have not considered the "Up and Down"

method of stochastic approximation.) The present area of investigation has similarities with sequential estimation, but also some important differences. For the estimation problem only the "stopping rule" need be considered, for no independent variable is manipulated.

The Problem

Assume that the experimenter's goal is that the value of a particular population mean is to be changed from its present value, δ, to a different value, α. For example, a population of adults may on the average score $\delta = 100$ on a particular standardized reading test and the goal is to increase that average to $\alpha = 116$. The experimenter has in mind a treatment variable (say, number of hours of individual tutoring) which he knows can affect the average reading score, but the exact nature of the relationship between reading score and tutoring is unknown. In other words, the "end" is known, but not the specific "means", and therefore, the appropriate value of the independent variable, or treatment condition, must be found. More formally, the expected value of the reading score is a function of the independent variable, $E(y) = f(x)$, and the experimenter wishes to determine that specific value of the independent variable, $x = \theta$, for which $E(y) = \alpha$, or $E(y(x = \theta)) = \alpha$. For present purposes it is assumed that if $x > \theta$ then $E(y(x)) > \alpha$, and if $x < \theta$ then $E(y(x)) < \alpha$. Given this situation the experimenter can select an initial value x_1 and thereafter choose the value of the independent variable as $x_{n+1} = x_n - a_n(y_n(x_n) - \alpha)$. The a_n are selected to have several characteristics, the most intuitively important of which is that $a_n \to 0$ as $n \to \infty$ "at a suitable rate". One possible definition is $a_n = \frac{1}{n}$. If appropriate a_n are chosen, such as $\frac{1}{n}$, Robbins and Monro (1951) proved $n \to \infty$, $x_n \to \theta$. The experimenter, of course, must have some feel for the

speed of convergence, and how this convergence is affected by the choice of x_1, the relationship between $E(y(x))$ and x, and the density of $y(x)$. He also must have some idea of when to stop the experimentation. Most of the results in the literature to date, however, are asymptotic in nature, with relatively little work being done on stopping rules (Chernoff (1975) offers no citations, but see Farrell (1962)). The literature, as it appears to us, provides little if any practical guidance for the experimenter.



Methods

Initial results were obtained with an interactive empirical approach using computer simulation techniques on a time-shared CDC Cyber 74. Many computer runs were made as the researchers sought to understand the importance of the numerous parameters which can be considered. Following this first phase of computer runs, during which all the values produced from a single sequential experiment were often observed, more traditional Monte Carlo experiments were performed, replicating the experiments a number of times to obtain estimates of how the procedures operate "in the long run". In summary, the approach used combined both an interactive search during which the researchers observed the behavior of various functional relationships during a single replication and more traditional "fixed" type of experiments in which a number of replications of an experimental situation were made to obtain stable estimators.

All pseudo random numbers were obtained from either NORMAL or RAN3F which are a normal (N) random number generator and a uniform (U) random number generator, respectively. One thousand random numbers were generated per each call of these routines and following generation they were immediately permuted by an independent randomization procedure using the program PERMUTE. All routines are maintained by the University of Minnesota Computer Center.

Design

The model for the random variable y was $y = \mu_{y.x} + \epsilon$, where $\mu_{y.x} = E(y(x)) = f(x) = \beta_0 x^0 + \beta_1 x^1 + \beta_2 x^2 + \beta_3 x^3$ and where ϵ is independently identically distributed either as $N(0,\sigma^2_{y.x})$ or $U(0,\sigma^2_{y.x})$. Following the interactive search in which many parametric specifications and stopping rules were studied, certain choices of parameters and rules were made for the more standard type of Monte Carlo investigation. These included:

1. Four definitions of a_n. They were $a_n = \frac{1}{n}, \frac{1}{n\beta'_\theta}, \frac{1}{n\hat{\beta}}, \frac{s^k}{n}$, where β'_θ is the first derivative of f(x) evaluated at θ, $\hat{\beta}$ is the usual slope estimator, and $s^k = |\sum_{i=j}^{n} z_i|^k$, $k = \text{INT}[s/\frac{1}{2}c]$, where $c > 0$ and even, $j = \max(1, n - c + 1)$ and $z_n = 1$ if $y(x_n) < \alpha$, or $z_n = -1$ if $y(x_n) > \alpha$. (INT means "integer part of.") For the procedure employing $\hat{\beta}$, $a_n = \frac{1}{n}$ for $n < 20$ and $\frac{1}{n\hat{\beta}}$ otherwise. In the definition $a_n = \frac{s^k}{n}$ a "finite memory" is introduced into the approximation process, and successively positive or negative values of $y_n(x_n) - \alpha$ cause larger adjustments x_{n+1} than is the case with the other definitions. Both s^k and $\hat{\beta}$ are random variables and this results in a variant of the Robbins-Monro procedure in that it assumes the a_n to be "a fixed sequence of positive constants."

2. Two stopping rules. They were

R1: Stop if $n \geq 20$ and if α contained in $\hat{\mu}_{y.x} \pm \hat{\sigma}_{y.x} \left(\frac{1}{n} + \frac{(x_n - \bar{x}_n)^2}{(\Sigma(x_i - \bar{x}_n)^2} \right)^{\frac{1}{2}} t_{p/2}$

(where $\hat{\mu}_{y.x} = \hat{\delta} + \hat{\beta} x_n$) or $n = 200$.

R2: Stop if $n \geq 20$ and, considering the last 20 values of z, if $\Sigma z = 9$, 10, or 11, and the number of "runs" is 9, 10, 11, 12, or 13, or if $n = 200$.

3. Three sets of β_0, β_1, β_2, β_3. They were [100, .14142, 0, 0] [100, .34641, 0, 0], and [100, .12686, .0058512, -.000023767].

4. Two conditional variances. They were $\sigma^2_{y.x} = 100$ and 25.

Most "final" experiments were based on 500 replications.[1] Based on these replications, the mean and variance were computed for $(x_n - \theta)$ at $n = 30$, 50, 100 steps and for both rules, R1 and R2. Additionally, for both rules the mean and variance of the number of steps needed to stop were also computed.

$a_n = \frac{1}{n}$ was included in the experiment because it was suggested in Robbins and Monro's original paper. Hodges and Lehmann provide results on $a_n = \frac{1}{n\beta_1}$ when the regression is in fact linear, and it has certain optimal characteristics and therefore was included as a basis for comparisons. In discussing the preceding work, Chernoff (1975) remarked that "In the stochastic approximation case using sequences $a_n = \frac{c}{n}$, there is no prior knowledge of θ to insure that $c = \beta^{-1}$. However, as data

accumulate one would hopefully obtain a satisfactory estimate of β providing the successive x_n are not too close to each other (p. 70)."[2] We interpreted these comments to mean that when one has "sufficient" information one would estimate β_1 using the least squares estimator, $\hat{\beta}$. Initial results demonstrated that the instability of $\hat{\beta}$ for small n caused erratic adjustments and poor convergence. This lead to the procedure $a_n = \frac{1}{n}$ for n = 1, . . ., 19 and $\frac{1}{n\hat{\beta}}$ thereafter. The $a_n = \frac{s^k}{n}$ procedure was developed during the interactive part of the present research. It seemed reasonable to specify an adjustment procedure which would make larger adjustments if $E(y_n(x_n)) - \alpha$ were judged to be large. Considering only the sign of $y_n(x_n) - \alpha$ and taking c = 4, for the patterns (+ + + +) or (- - - -), $s^k = 16$, for patterns like (+ - + +) or (- - - +) $s^k = 2$, and for patterns with two pluses and two minuses, $s^k = 1$. This type of adjustment assumes that the error distributions are symmetric so that the probability of a plus at $x_n = \theta$ is $\frac{1}{2}$. During the interactive phases of this research, c = 4 and c = 10 were found to work well.

Stopping rule R1 employs the standard confidence interval for estimating $\mu_{y.x}$. This seemed a reasonable approach to consider, especially when $a_n = \frac{1}{n\hat{\beta}}$ is employed. The confidence coefficient, p, used was .60. This value was chosen during the interactive phase on the basis of performance.

Stopping rule R2 comes from reasoning similar to that used in developing the s^k procedure. At $x_n = \theta$, for symmetric error distributions

the sign of $y_{\ln}(x_n) - a$ would be independently distributed as a Bernoulli variable with parameter $\frac{1}{2}$. R2 essentially tests two hypotheses, one concerning "randomness" and the other that the proportion of "pluses" is $\frac{1}{2}$.

Results

Results of the runs are presented in Tables 1-8. The mean and variance of the bias $(x_n - \theta)$, are reported for the conditions studied as well as the mean and variance of the "number of steps to stop" for R1 and R2. Average squares bias, $(x_n - \theta)^2$ is not reported, but it can be easily obtained by squaring the average bias and adding this to the variance of the bias (i.e. $E(x_n - \theta)^2 = V(x_n - \theta) + (E(x_n - \theta))^2$).

The method employing $\frac{1}{n\delta'_\theta}$ was superior to the other methods, but since δ'_θ would seldom be known, results associated with δ'_θ will be of greatest utility as "benchmarks". It is clear from the results that generally, the procedure $a_n = \frac{1}{n}$ had the poorest performance. In situations where little or no prior information is available about the relationship between the independent and dependent variables, the results would lead us to use $a_n = \frac{s^k}{n}$ with $c = 4$. This typically does as well as or better than the other procedures not employing δ'_θ. It is markedly better when there is a weak relationship between the independent and dependent variables and a poor start is made (see columns 1 and 2 of Tables 1-7). In an attempt to determine the behavior of $a_n = \frac{s^k}{n}$ when a "good start" is made, the experiments reported in Table 8 with $x_1 = \theta$ were carried out. We believe the procedure did reasonably well under these circumstances.

If more information is available, one might profitably choose one of the procedures studied. Neither R1 nor R2 is uniformally better with respect to bias and number of steps to stop. There also appears to be at least some interaction with the definition of a_n, and this complicates matters in a few instances. Here we can only recommend that one make a best guess about conditions and use that stopping rule which would be best.

Educational Significance

One potential area of application for stochastic approximation is that of formative evaluation. Stochastic approximation can suggest values of the independent variable which would attain programmatic goals, and this information could be fed to persons directly involved in program development. Within the framework developed by Sanders and Cunningham (1974), stochastic approximation could provide "external information" for "formative interim evaluation activities". When a summative evaluation is planned, perhaps using one of the more standard experimental designs, design points can be chosen in the region suggested through sequential experimentation, thereby increasing the likelihood that the program will demonstrate its effectiveness.

In general, stochastic approximation would appear to be a useful technique in any area where individuals have a goal firmly in mind but lack sufficient knowledge of the independent variable to design an efficient, more traditional experiment. Education is goal oriented, and information about how to achieve a goal is often more important than, say, information about the exact nature of the relationship between an independent and dependent variable. Stochastic approximation can provide useful information about an independent variable, even when its defined over a broad range of values, while requiring relatively few subjects for its implementation.

Table 1

Mean and Variance of the Bias $(x_n - \theta)$ at 30 Steps Where $x_1 = 4$[†]

a_n	$\sigma^2_{y.x}$	$[\beta_0,\beta_1,\beta_2,\beta_3]$					
		[100,.14142,0,0]		[100,.34641,0,0)		[100,.12686,.0058512, −.000023767]	
		$\epsilon\sim N(0,\sigma^2_{y.x})$	$\epsilon\sim U(0,\sigma^2_{y.x})$	$\epsilon\sim N(0,\sigma^2_{y.x})$	$\epsilon\sim U(0,\sigma^2_{y.x})$	$\epsilon\sim N(0,\sigma^2_{y.x})$	$\epsilon\sim U(0,\sigma^2_{y.x})$
$\frac{1}{n\beta'_\theta}$	100	-.49	.38	-.13	.20	-.96	-.79
		160.93	177.75	27.18	29.84	12.85	14.54
	25	-.19	.24	-.08	.10	-.53	-.48
		40.85	44.77	6.81	7.46	3.39	3.81
$\frac{1}{n}$	100	-61.16	-61.11	-9.73	-9.25	-6.57	-6.68
		72.00	81.11	27.65	32.02	20.82	22.91
	25	-61.27	-60.96	-9.46	-9.62	-6.48	-6.36
		18.89	20.48	7.69	7.62	4.74	5.59
$\frac{1}{n\hat{\beta}}$	100	-44.86	-41.86	-5.65	-4.55	-3.14	-2.36
		1436.76	1385.58	140.41	180.92	108.84	188.85
	25	-36.23	-35.41	-5.72	-5.57	-3.07	-3.10
		490.86	577.45	65.93	51.08	43.88	50.68
$\frac{s^k}{n}(c=4)$	100	-1.92	-3.58	.42	.02	-.18	-.19
		298.62	452.01	72.73	97.26	44.55	61.85
	25	-.88	-.98	.44	.19	.18	-.32
		71.44	96.15	18.46	22.93	10.77	14.73
$\frac{s^k}{n}(c=10)$	100	-10.85	-17.03	-.81	-1.93	-.06	-.89
		835.47	764.57	110.47	87.05	47.69	49.23
	25	.37	-1.98	-.67	-.75	-.10	.26
		378.17	419.80	24.05	25.86	24.08	11.04

[†]For Tables 1-8 the upper number in each cell is the mean and the lower number is the variance.

Table 2

Mean and Variance of the Bias $(x_n - \theta)$ at 50 Steps Where $x_1 = 4$

a_n	$\sigma^2_{y.x}$	$[\beta_0,\beta_1,\beta_2,\beta_3]$					
		[100,.14142,0,0]		[100,.34641,0,0]		[100,.12686,.0058512, −.000023767]	
		$\varepsilon\sim N(0,\sigma^2_{y.x})$	$\varepsilon\sim U(0,\sigma^2_{y.x})$	$\varepsilon\sim N(0,\sigma^2_{y.x})$	$\varepsilon\sim U(0,\sigma^2_{y.x})$	$\varepsilon\sim N(0,\sigma^2_{y.x})$	$\varepsilon\sim U(0,\sigma^2_{y.x})$
$\frac{1}{n\beta'_\theta}$	100	.30	.34	.17	.16	-.43	-.58
		93.84	109.69	15.75	18.35	7.21	8.67
	25	.18	.20	.08	.08	-.28	-.25
		23.66	27.54	3.94	4.59	1.96	2.25
$\frac{1}{n}$	100	-56.87	-56.83	-8.02	-7.69	-5.04	-5.17
		62.61	69.82	20.59	23.88	13.69	15.49
	25	-56.92	-56.61	-7.92	-8.04	-4.98	-4.88
		16.49	17.82	5.44	5.88	3.12	3.67
$\frac{1}{n\hat{\beta}}$	100	-20.43	-22.29	-1.38	-1.51	-.63	-.67
		1036.14	994.74	158.51	104.35	33.50	67.36
	25	-20.56	-20.91	-2.97	-3.20	-1.47	-1.43
		231.29	236.72	25.52	18.06	12.84	13.58
$\frac{s^k}{n}(c=4)$	100	-1.48	-1.99	.07	-.06	.26	-.34
		157.02	262.19	39.42	50.39	25.33	33.99
	25	-.46	-.47	.27	.11	.04	.07
		40.04	58.47	9.83	13.56	6.11	8.22
$\frac{s^k}{n}(c=10)$	100	-7.93	-12.07	-.38	-1.36	-.12	-.31
		476.52	513.41	54.41	46.98	22.12	22.28
	25	.33	-1.81	-.10	-.02	.20	.16
		169.92	229.20	12.92	12.72	6.35	6.72

Table 3

Mean and Variance of the Bias ($x_n - \theta$) at 100 Steps Where $x_1 = 4$

			$[\beta_0, \beta_1, \beta_2, \beta_3]$			
		[100,.14142,0,0]		[100,.34641,0,0]	[100,.12686,.0058512, -.000023767]	
a_n	$\sigma^2_{y.x}$	$\varepsilon \sim N(0,\sigma^2_{y.x})$	$\varepsilon \sim U(0,\sigma^2_{y.x})$	$\varepsilon \sim N(0,\sigma^2_{y.x})$	$\varepsilon \sim U(0,\sigma^2_{y.x})$	$\varepsilon \sim N(0,\sigma^2_{y.x})$	$\varepsilon \sim U(0,\sigma^2_{y.x})$
$\frac{1}{n\beta'_\theta}$	100	.35 48.75	.00 58.39	.16 8.14	.01 9.75	-.16 3.56	-.27 4.65
	25	.19 12.20	.01 14.63	.08 2.03	.00 2.44	-.11 .95	-.16 1.08
$\frac{1}{n}$	100	-51.56 53.31	-51.46 57.95	-6.31 13.19	-6.05 15.68	-3.53 7.21	-3.63 8.42
	25	-51.56 13.64	-51.36 14.56	-6.20 3.50	-6.28 3.83	-3.48 1.71	-3.42 2.07
$\frac{1}{n\hat{\beta}}$	100	-8.62 333.15	-9.93 308.16	-.43 44.44	-.40 34.35	-.16 8.90	-.31 18.12
	25	-10.44 74.54	-10.67 71.06	-1.39 7.19	-1.53 5.60	-.50 3.08	-.58 3.30
$\frac{s^k}{n}(c=4)$	100	-.28 78.34	-1.03 115.32	.25 21.10	.15 27.90	.09 11.89	.02 15.02
	25	-.11 19.31	-.27 28.79	.05 5.08	-.04 6.84	-.05 3.10	.02 4.06
$\frac{s^k}{n}(c=10)$	100	-5.75 227.50	-7.97 299.75	-.35 19.40	-.61 35.03	.04 11.50	-.29 14.37
	25	-.26 71.40	-1.61 113.45	.00 5.18	-.06 8.52	.12 2.98	-.01 3.15

Table 4

Mean and Variance of the Bias ($\bar{x}_n - \theta$) When

Stopped with Parametric Rule (R1) Where $x_1 = 4$

$[\beta_0, \beta_1, \beta_2, \beta_3]$

a_n	$\sigma^2_{y.x}$	$\varepsilon \sim N(0, \sigma^2_{y.x})$	$\varepsilon \sim U(0, \sigma^2_{y.x})$	$\varepsilon \sim N(0, \sigma^2_{y.x})$	$\varepsilon \sim U(0, \sigma^2_{y.x})$	$\varepsilon \sim N(0, \sigma^2_{y.x})$	$\varepsilon \sim U(0, \sigma^2_{y.x})$
		[100,.14142,0,0]		[100,.34641,0,0]		[100,.12686,.0058512, -.000023767]	
$\frac{1}{n\beta'_\theta}$	100	-.70	.66	-.17	.33	-1.07	-1.09
		237.97	265.39	40.34	44.80	18.41	21.33
	25	-.25	.40	-.10	.16	-.49	-.55
		60.46	67.14	10.08	11.18	4.75	5.50
$\frac{1}{n}$	100	-47.45 *	-46.62 *	-7.82	-7.49	-4.42	-4.54
		61.09	49.33	23.91	24.05	13.57	14.73
	25	-46.58 *	-46.76 *	-4.94 *	-5.04 *	-2.60	-2.54 *
		11.45	12.53	2.87	2.70	1.38	1.26
$\frac{1}{n\hat{\beta}}$	100	-14.39	-12.61	-5.28	-5.11	-2.70	-2.56
		924.61	1000.61	78.38	96.76	44.00	47.93
	25	-5.16	-5.46	-2.04	-1.92	-.68	-.74
		112.84	71.87	13.92	12.36	5.93	3.91
$\frac{s^k}{n}(c=4)$	100	-2.36	-4.32	.06	.41	-.80	-1.15
		352.92	460.81	59.17	67.44	28.71	32.94
	25	-.89	-1.20	.10	.22	-.32	-.65
		91.96	110.92	14.93	15.90	7.28	7.61
$\frac{s^k}{n}(c=10)$	100	-13.25	-14.75	-1.91	-2.41	-.42	-.71
		498.27	502.03	77.19	69.56	32.00	33.87
	25	-1.88	-1.79	-.48	-.64	.42	.39
		129.16	114.31	14.53	17.57	8.27	8.20

*The estimates in these cells are based on 100 replications instead of 500.

Table 5

Mean and Variance of the Number of Steps

When Stopped with Parametric Rule (R1) Where $x_1 = 4$

		$[\beta_0,\beta_1,\beta_2,\beta_3]$					
		[100,.14142,0,0]		[100,.34641,0,0]		[100,.12686,.0058512, -.000023767]	
a_n	$\sigma^2_{y.x}$	$\varepsilon\sim N(0,\sigma^2_{y.x})$	$\varepsilon\sim U(0,\sigma^2_{y.x})$	$\varepsilon\sim N(0,\sigma^2_{y.x})$	$\varepsilon\sim U(0,\sigma^2_{y.x})$	$\varepsilon\sim N(0,\sigma^2_{y.x})$	$\varepsilon\sim U(0,\sigma^2_{y.x})$
$\frac{1}{n\beta'_\theta}$	100	20.18	20.21	20.13	20.18	23.13	22.23
		2.37	1.79	1.30	1.54	338.95	194.61
	25	20.17	20.13	20.18	20.14	26.28	24.96
		2.07	1.08	2.33	1.09	622.57	464.60
$\frac{1}{n}$	100	196.44	198.24	87.64	85.30	92.26	95.61
		627.30 *	309.76 *	7087.82	7003.51	7320.44	7372.14
	25	200.00	200.00	191.49 *	196.05 *	191.22	194.62 *
		0.00 *	0.00 *	1393.06	572.69	1416.44	622.76
$\frac{1}{n\hat{\beta}}$	100	89.66	93.18	28.76	30.11	28.78	29.03
		5019.12	5523.06	788.86	1132.37	695.83	716.53
	25	177.87	178.80	76.85	82.71	78.21	80.39
		2849.49	2840.74	2214.93	2414.49	1589.28	1468.92
$\frac{s^k}{n}(c=4)$	100	24.86	27.53	25.68	26.46	26.02	27.16
		338.35	610.14	210.89	262.41	165.28	228.26
	25	25.03	27.37	25.50	26.75	25.87	28.97
		477.47	711.50	155.27	213.57	137.34	405.82
$\frac{s^k}{n}(c=10)$	100	53.20	65.79	24.32	25.61	23.96	26.59
		3807.48	5113.46	399.95	492.67	250.35	654.13
	25	83.42	111.00	27.24	29.36	27.01	26.12
		6015.15	6715.59	775.31	1136.49	777.34	646.42

*The estimates in these cells are based on 100 replications instead of 500.

Table 6

Mean and Variance of the Bias $(x_n - \theta)$ When

Stopped with Nonparametric Rule (R2) Where $x_1 = 4$

		$[\beta_0,\beta_1,\beta_2,\beta_3]$					
		[100,.14142,0,0]		[100,.34641,0,0]		[100,.12686,.0058512, -.000023767]	
a_n	$\sigma^2_{y.x}$	$\varepsilon\sim N(0,\sigma^2_{y.x})$	$\varepsilon\sim U(0,\sigma^2_{y.x})$	$\varepsilon\sim N(0,\sigma^2_{y.x})$	$\varepsilon\sim U(0,\sigma^2_{y.x})$	$\varepsilon\sim N(0,\sigma^2_{y.x})$	$\varepsilon\sim U(0,\sigma^2_{y.x})$
$\frac{1}{n\beta'_\theta}$	100	-.21	.05	-.01	.07	-.94	-.80
		173.52	179.81	29.40	29.87	12.59	14.97
	25	.00	.08	.00	.03	-.44	-.54
		44.00	44.88	7.34	7.48	3.42	4.09
$\frac{1}{n}$	100	-48.05	-50.04	-7.75	-8.02	-4.68	-5.52
		42.17	51.09	16.73	25.63	11.06	17.34
	25	-46.74	-46.54	-6.17	-7.01	-3.66	-4.10
		11.18	12.08	3.10	4.97	1.75	2.90
$\frac{1}{n\hat{\beta}}$	100	-9.77	-14.57	-2.86	-3.25	-1.55	-2.16
		526.20	631.61	54.60	84.38	18.48	30.35
	25	-9.70	-11.74	-2.53	-3.42	-1.48	-1.83
		56.80	82.45	8.40	11.26	4.10	5.39
$\frac{s^k}{n}(c=4)$	100	-.46	-2.30	.53	-.20	-.47	-.60
		249.88	453.42	77.44	106.10	46.65	65.27
	25	-.04	-.50	.49	.07	.12	-.11
		61.75	92.97	16.49	25.41	10.77	14.87
$\frac{s^k}{n}(c=10)$	100	-8.11	-12.36	-.34	-1.73	.29	-.25
		391.91	502.21	42.68	64.64	30.43	33.71
	25	.49	-1.48	.00	-.12	.61	.52
		138.40	215.96	10.91	16.19	6.55	7.37

Table 7

Mean and Variance of the Number of Steps

When Stopped with Nonparametric Rule (R2) Where $x_1 = 4$

		$[\beta_0,\beta_1,\beta_2,\beta_3]$					
		[100,.14142,0,0]		[100,.34641,0,0]		[100,.12686,.0058512, -.000023767]	
a_n	$\sigma^2_{y.x}$	$\varepsilon\sim N(0,\sigma^2_{y.x})$	$\varepsilon\sim U(0,\sigma^2_{y.x})$	$\varepsilon\sim N(0,\sigma^2_{y.x})$	$\varepsilon\sim U(0,\sigma^2_{y.x})$	$\varepsilon\sim N(0,\sigma^2_{y.x})$	$\varepsilon\sim U(0,\sigma^2_{y.x})$
$\frac{1}{n\beta'_\theta}$	100	32.42	31.79	32.40	31.86	35.06	33.45
		226.32	187.77	223.18	187.39	294.07	247.42
	25	32.60	31.98	32.60	31.98	34.81	33.31
		224.71	187.41	224.71	187.41	240.51	237.54
$\frac{1}{n}$	100	173.71	136.65	61.13	50.38	60.90	48.90
		2234.70	3301.21	1369.02	917.82	1359.13	756.51
	25	200.00	199.66	112.97	84.67	101.78	79.81
		0.00	34.43	2832.98	1962.52	2279.99	1652.54
$\frac{1}{n\hat{\beta}}$	100	84.30	70.84	51.21	43.98	47.11	44.06
		1188.04	999.99	1039.26	681.22	596.34	642.90
	25	110.54	94.07	66.81	57.02	59.28	53.61
		1751.84	1258.61	718.70	614.92	536.85	471.00
$\frac{s^k}{n}(c=4)$	100	39.48	36.76	31.71	30.31	30.70	30.02
		365.17	341.61	203.34	156.66	191.41	172.76
	25	36.85	36.51	35.16	31.93	31.08	30.98
		239.20	297.97	310.12	171.47	140.91	162.65
$\frac{s^k}{n}(c=10)$	100	58.92	54.05	39.62	36.75	35.52	35.71
		994.45	745.88	431.32	358.61	307.83	340.59
	25	64.06	59.23	39.79	37.49	37.07	34.87
		1192.78	953.26	407.40	324.22	364.19	315.87

Table 8

Means and Variances Where $x_1=\theta$ and $\varepsilon \sim N(0, \sigma^2_{y.x}=100)$

$[\beta_0,\beta_1,\beta_2,\beta_3]$ a_n	Bias $(x_n-\theta)$ at 30 steps	Bias $(x_n-\theta)$ at 50 steps	Bias $(x_n-\theta)$ at 100 steps	Bias $(x_n-\theta)$ when stopped with R1	No. of steps when stopped with R1	Bias $(x_n-\theta)$ when stopped with R2	No. of steps when stopped with R2
$\frac{1}{n\beta'_\theta}$	-.49	.30	.35	-.88	20.33	-.22	32.59
[100,.14142,0,0]	160.93	93.84	48.75	235.51	1.11	172.67	222.70
$\frac{s^k}{n}(c=4)$	.30	.10	.39	-.03	23.82	-.25	33.11
	234.71	148.69	77.78	243.27	80.88	232.17	289.16
$\frac{1}{n\beta'_\theta}$	-.13	.17	.16	-.26	20.30	-.01	32.57
[100,.34641,0,0]	27.18	15.75	8.14	39.86	1.07	29.24	219.56
$\frac{s^k}{n}(c=4)$	.25	.17	.08	-.06	25.45	.66	30.56
	72.55	36.14	19.10	63.29	176.93	77.57	203.33

References

Chernoff, H. Approaches in sequential design of experiments. in J. N. Srivastava (Ed.), A survey of statistical design and linear models. Amsterdam: North-Holland Publishing, 1975.

Farrell, R. H. Bounded length confidence intervals for the zero of a regression function. Annals of Mathematical Statistics, 1962, 33, 237-247.

Herzberg, A. M. & Cox, D. R. Recent work on the design of experiments: A bibliography and a review. Journal of the Royal Statistical Society, 1969, Series A, 132, 29-67.

Hodges, J. L. & Lehmann, E. L. Two approximations to the Robbins-Monro process. In J. Neyman (Ed.), Proceedings of the Third Berkley Symposium on Mathematical Statistics and Probability. Berkeley, California: University of California Press, 1956.

Lord, F. M. Robbins-Monro procedures for tailored testing. Educational and Psychological Measurement, 1971, 31, 3-31.

Lord, F. M. Tailored testing, An application of stochastic approximation. Journal of American Statistical Association, 1971, 66, 707-711.

Meyer, D. L. Response surface methodology in education and psychology. The Journal of Experimental Education, 1963, 31, 329-336.

Robbins, H. & Monro, S. A stochastic approximation method. Annals of Mathematical Statistics, 1951, 22, 400-407.

Sanders, J. R. & Cunningham, D. J. Formative evaluation: Selecting techniques and procedures. In G. D. Borich (Ed.), Evaluating educational programs and products. Englewood Cliff, New Jersey, Educational Technology Publications, 1974.

Venter, J. H. An extension of the Robbins-Monro procedure. Annals of Mathematical Statistics, 1967, 38, 181-190.

Wetherill, G. B. Sequential methods in statistics. London: Chapman and Hall, 1975.

Footnotes

[1]The original plan was to have 500 replications for each set of conditions, however, given the value of x_1 used here, $a_n = \frac{1}{n}$ converged slowly and for some conditions the rule "stop if n = 200" was used for virtually every replication. We decided to use only 100 replications in these instances, and those runs are noted in the tables.

[2]The "c" in this quote is not defined in the same way as the "c" in the definition of $a_n = \frac{s^k}{n}$.